USING MOLECULAR MODELING TO DESIGN SEMI-FLEXIBLE POLYMERIC MATERIALS

A Dissertation
Presented to the Faculty of the Graduate School
of Cornell University
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

By

Christian Nowak

August 2019

USING MOLECULAR MODELING TO DESIGN SEMI-FLEXIBLE POLYMERIC MATERIALS

Christian Nowak, Ph.D.

Cornell University 2019

Using Molecular Dynamics, we simulate detailed All-Atom models of (at the time) unsynthesized polythiophene derivatives with oligoethylene glycol side chains under the stimulus of an applied electric field to evaluate their ionic conduction properties. The originally proposed chemistry (P3MEET) has an oxygen atom in the side chain bonded directly to the polythiophene backbone which is identified as detrimental to the ionic mobility as the mobility of that oxygen is hampered by the covalent bond to the backbone. The simple insertion of a single methylene group between this oxygen and the backbone (P3MEEMT) leads to a near five-fold improvement in the ionic conductivity as observed in our simulations. Results from these studies informed experimental collaborators leading to experimental evidence verifying the accuracy of our initial predictions.

We continue by examining more chemistries which take the three oxygens in the side chain of P3MEEMT and goes through every permutation of replacing any number of the oxygens with a methylene group. The results of this study show that the main factors that are in competition are the tendency to reduce the oxygen concentration to reduce the average number of oxygens that coordinate a given ion and the percolation of solvation sites. While a reduced number of coordinating oxygens is beneficial as it becomes easier to escape that

trap, it also correlates to an increased distance between solvation sites. This weakened percolation of solvation sites increases the energetic barrier as the ion needs to make a longer jump between sites. This competition is observed only in the amorphous domain because the crystalline domain has a regular arrangement of side chains which automatically creates a percolated network of nearby solvation sites, masking this issue.

We use generic coarse-grained (CG) models which can represent many different polymers to study networks of semi-flexible block copolymer chains (D-LCE) which exhibit a saw-tooth tensile response when subjected to a strain field. A simulated synthesis shows the synthetic viability of these networks which leads to an exploration of the block architecture and segregation strength of the chains to elucidate new behaviors or improved properties in the deformation mechanics. These studies show that this class of materials has promise and should be studied. To aid experimentalists in finding a chemistry for these systems, we have developed a methodology to begin the process of inverse coarse graining (ICG) which aims to find a molecule whose physics/behavior fits well onto the physics/behavior of a CG model of interest. Outside of our studies, generic CG models which don't represent a specific chemistry are used extensively to scan design spaces in search of novel behaviors/properties. As such, the ICG process is critical if computer simulations using generic CG models are to inform experimentalists.

We also present a method for calculating the free energy of polymer self-assembly into different mesophases. A guiding field approach is used to take a system of athermal chains in the isotropic phase and assembly them in into the target mesophase which circumvents the often problematic first-order transition that is seen in polymer self-assembly. Using this

approach the phase behavior or rod-coil block copolymers is studied as well as the statistical mechanical origins for differences in the phase behavior of two nearly identical molecules who only vary in the their architecture.

Biographical Sketch

Chris was born in Bristol, CT, USA in 1992. With a penchant for Maths and Curiosity for the unconventional he acquired a Bachelor's of Science in Chemical Engineering from Rensselaer Polytechnic Institute in May 2014. Thanks to healthy competition with a roommate, Chris joined the department of Chemical and Biomolecular engineering at Cornell University as a Ph.D. student in August 2014. Officially joining the Escobedo group in January 2015, he has worked on a broad range of problems relating to simulation of primarily polymeric materials with a focus on computationally-led materials design. When not working he enjoys making noises with stringed instruments, climbing up things, and picking heavy things up.

Dedication & Acknowledgments

This thesis is dedicated to my parents, Anna and Wlodzimierz. I thank them for all the sacrifices they've made to put me in a position where I could pursue whatever I set my mind to. The furrier member of the family, Caramel, also deserves acknowledgment for keeping the family together even through the tougher moments of the last 5 years and not peeing on the carpet too many times. My brother from another mother, Mike, is also thanked for many reasons.

I thank my mentor in this journey, Fernando, for always considering the ideas I bring, knowing when to steer me in the right direction and knowing when to allow me to explore whatever caught my fancy at the time. This positive mentor-mentee relationship has been integral to my success and my enjoyment of this whole process. What also aided in success is colleagues such as Endian, Mayank, Unmukt, and Poornima who are always willing to give feedback on my ideas and laugh at my terrible jokes. Even more instrumental is Gabe and Alex initially planting the idea of seeking graduate study, without which I would not even be writing this. I came into Cornell woefully unprepared, without any intention of doing simulations so the fact that this experience has turned out so well is something I am truly grateful for.

Finally, I'd like to thank Hannah, Lindsay, and Ray for innumerable reasons. Thank you all.

Table of Contents

List of Figures

List of Tables

1 Introduction

As the search for new materials with improved or novel properties continues, there is increasing stress placed on accelerating this process, as emphasized by large organizations such as the Materials Genome Initiative.[39,94] The primary strategy has been to increase collaboration between experimentalists and simulation groups to leverage the advantages of both approaches. The advantages of computer simulations include: (i) the ability to assess chemistries that have yet to be synthesized by experimentalists, (ii) atomic level information offering key insights into the physical mechanisms, and (iii) the ability to impose unphysical constraints when needed to attain desirable information. Any of these strengths can be emphasized by an appropriate choice of method ranging from larger scale information by using field-theoretic models such as Self-Consistent Field Theory, to atomistic scale information with Molecular Dynamics (MD) and Monte Carlo simulations, to electron orbital scale information with Density Functional Theory. A class of molecules which have benefitted greatly from computer simulations are macromolecules (polymers) which are difficult to synthesize and therefore experimentalists can benefit from the knowledge gleamed from simulations, providing guidance as to what avenues are more promising to explore as opposed to searching randomly.

Under the broad umbrella of polymeric materials, those which possess semi-flexibility are of interest. One subset of these materials, dual-conducting polythiophenes with oligoethylene glycol side chains is of interest because that which causes the semi-flexibility, the delocalized π electron cloud, also confers electronic conduction properties[189]. Polymers are usually

thought of as insulators so the ability to combine this functionality with more the common ionic conductivity leads to interesting applications ranging from bioelectronics[104,123,154,173] to energy applications including batteries,[40,180] pseudo capacitors,[126] electrochromic windows,[139] dye-sensitized solar cells[167], and mechanical actuators.[77] The "soft" nature of conjugated polymers means that they will not damage any biological entity they are interfaced with, enabling a high quality interface, making them particularly attractive for bioelectronics applications[90,173].

While electronic conduction in conjugated polymers has been well studied, ionic transport has been explored only on a limited basis. Detailed investigations on the influence of polymer chemistry on ion uptake, polymer morphology and ionic transport are lacking so we aim to use molecular simulations to lead the design of new conjugated polymers, in particular polythiophene derivatives with oligoethylene glycol side chains. The advantage of simulations here is that many chemistries which are not yet synthesized can be studied quickly in parallel to give guidance to experimental collaborators when they seek to synthesize a target material.

Using MD simulations, we study the ionic conductivity characteristics of Li^+ ions in these systems for both the amorphous and crystalline domain as these materials show complex semi-crystalline mesophases. We study a broad range of side chain chemistries to understand how the concentration and arrangement of strongly partially charged groups in the side chain affect the physics governing ionic conduction. We summarize these results using design principals which can vary depending on which of the two phases is more present in the semi-crystalline mesophase.

The second semi-flexible polymeric material we are interested in are liquid crystalline elastomers (LCE). LCE are cross-linked networks composed of semi-flexible polymer chains. Here the semi-flexible nature of the chains is of interest because it leads to orientational alignment when they are uniaxially deformed at appropriate temperatures and densities[44,188], or when exposed to other stimuli such as electric fields[34], temperature[204], or light[124]. This ordering is another method of controlling molecular organization besides the more common incorporation of chemically incompatible blocks in the chains[63,234]. Additionally, the semi-flexible nature confers anisotropic shape changes of the network when the chains orientationally order due to the prolate chain shape, a property central to artificial muscles[125,222] and high-strain actuators[218,232]. This macroscopic shape change highlights the coupling between material shape and chain orientation.

LCEs also display nonlinear elasticity, as seen in actin cytoskeleton networks[95,225], and spider dragline silk[110,228]. One reason for their nonlinear elasticity is the loss of entropic elasticity due to the existence of hairpins and other similar structures. For a chain hairpin, the entropy for unequal leg length is the same as that with equal leg length due to minimal transverse chain motion[3]. As such, the chain end-to-end distance does not correlate with the force, as only when hairpins unfold does entropy change.

Simulations of LCEs with diamond connectivity and semiflexible block copolymer chains show a non-linear deformation behavior[4,17] stemming from network chain hairpin unfold during deformation. New segregated domains of cross-links form to relax the stress and

allow further extension. Due to the immiscibility between the two blocks, there is an increase in interfacial energy when forming new domains, giving rise to a very pronounced "saw-tooth" type of tensile behavior. We aim to answer several questions with respect to D-LCE. First, how does a network with defects and entanglements affect the tensile response? The initial studies focused on defect-free and entanglement-free networks ("perfect" networks) so the tensile response of non-perfect network topologies needs to be understood. Second, if our studies indeed show that non-perfect network topologies can reproduce the saw-tooth tensile response, can altering the chain design, block architecture, or immiscibility of blocks lead to new or improved tensile responses?

The final piece of our studies is the development of computational tools which can aid in the materials discovery process. One complication of the materials discovery process through simulations is that often "identityless" models, i.e. models where the specific chemistry it represents is unknown, are used to scan large design spaces and make predictions. While instructive for understanding the physics of the problem at hand, it doesn't provide a chemistry for chemists to target, which is a primary goal for simulations. To facilitate this, we seek to develop a methodology which begins to tackle this ill-defined problem.

The other complication which can be seen in simulations which seek to map the self-assembly phase boundaries and the coexistence of two different self-assembled mesophases is observed. This is seen when the difference in free energy of self-assembly between the two mesophases is similar, leading to the slow kinetics of going from a metastable system to a marginally more stable system. The only absolute method for

discriminating between the stability of mesophases phases is the calculation of the Free Energy of self-assembly for both phases. This process is tricky as phase transitions are usually first-order, which incurs some spontaneous change in free energy which often cannot be captured. We seek to apply a field based approach used in earlier work[157] which transforms the first-order transition into a second-order transition, eliminating the complications found in non-field based approaches. Using this method we seek to understand the self-assembly of block copolymers with semi-flexible and flexible blocks[151] as well as the reason for why a simple change in the architecture in a side chain of bolaamphiphilic molecules[201] leads to profound differences in the self-assembly behavior.

2 Influence of Side-Chain Chemistry on Structure and Ionic Conduction Characteristics of Polythiophene Derivatives: A Computational and Experimental Study

2.1 Introduction

Due to their solution processability, mechanical flexibility and the fact that facile molecular design/synthesis strategies can be readily employed to tailor their optical and electronic properties, conjugated polymers continue to be of interest in scientifically and commercially diverse applications.[47,84,96,158,185,189] Traditionally, the electronic transport functionality of conjugated polymers have long played a vital role in the advancement of organic electronics applications such as xerography,[22] organic light-emitting diodes (OLEDs),[66,140] organic thin-film transistors (TFTs)[189] and organic photovoltaics (OPVs).[76,114] More recently, conjugated polymers have also made their way into many novel electrochemical applications that require simultaneous electronic and ionic conducting capability raging from bioelectronics[104,123,154,173] to energy applications including batteries,[40,180] pseudo capacitors,[126] electrochromic windows,[139] dye-sensitized solar cells[167], and mechanical actuators.[77] In addition to the inherent advantages over their inorganic counterparts, the "soft" nature of conjugated polymers and other organic materials enables the high quality interface between biological entities and electronics, rendering them particularly useful for bioelectronics applications.[91,173]

The key performance metrics of devices based on mixed ionic/electronic polymeric conductors such as mechanical, electrical or biological responses rely on electronic and ionic transport, which are strongly bound to the morphology of the polymers. Therefore, designing high performance mixed conducting polymers requires an understanding of the

interplay between polymer chemistry, processing, and morphology on the ionic and electronic conduction characteristics.[69] In contrast to inorganic semiconductors where the atoms are covalently bonded, the weak intermolecular bonding between macromolecular chains in conjugated polymers results in numerous degrees of conformational freedom, giving rise to their complicated microstructures. The typical structure of conjugated polymers is semicrystalline, consisting of crystalline/aggregate domains embedded in a disordered amorphous matrix. Past work has shown convincing evidence suggesting that electronic carriers traverse mainly through the crystalline domains and only occasionally transport through amorphous domains via bridging tie-chains.[45,50,108,135,235] Ionic transport in conjugated polymers, however, has been explored only on a limited basis. Most of the ion transport studies on conjugated polymers have been performed on poly(3,4-ethylenedioxythiophene) doped with poly (styrene sulfonate) (PEDOT: PSS) which has a rather complicated structure, preventing its use as a model system for structure-function relationship studies.[89,172,196] Despite recent work focusing on developing "non-PEDOT" conjugated polymers having polar side-chains and ion-conducting functionality,[68,143,162] detailed investigations on the influence of polymer chemistry on ion uptake, polymer morphology and ionic transport are still lacking.

To this end, in this work we report a systematic investigation on influence of ion uptake on structure and ion conducting behavior of two polythiophene derivatives bearing different oligoethylene glycol side-chains. Using molecular dynamics (MD) simulation, we show that P3MEET, a polythiophene derivative bearing oligoethylene glycol side-chains with an oxygen bonded directly to the thiophene rings, has lower segmental mobility compared to its analog P3MEEMT that have a methyl spacer between the oxygen and thiophene rings.

Consequently, the Li^+ ions conductivity in amorphous P3MEEMT is predicted to be approximately a factor of 5 higher than in amorphous P3MEET. Structural characterization of P3MEET and P3MEEMT thin film upon blending with lithium bis(trifluoromethanesulfonyl)imide (LiTFSI) indicates that LiTFSI infiltrates into both crystalline and amorphous domains at low LiTFSI concentration but preferably resides in amorphous domain at high LiTFSI concentration. Using a combination of simulation and experiment results, we infer that ionic transport in both P3MEET and P3MEEMT occurs in amorphous domain. Moreover, in excellent agreement with our MD simulation prediction, the experimental ionic conductivity measured by electrochemical impedance spectroscopy (EIS) in P3MEEMT is always higher than in P3MEET at all temperatures and concentrations studied. Our results highlight the importance of side-chain chemistry in dictating ionic transport of conjugated polymers.

2.2 Results & Discussion

In the first phase of our investigation, we employ molecular dynamics (MD) simulations to search for relevant chemistries in order to guide complementary synthesis and characterization efforts. We chose to use polythiophene derivatives due to the vast literature available on their structure and charge transport models.[14,37,81,98,108,146] Additionally, the morphological structure of polythiophene can be investigated with common techniques such as X-ray scattering or optical methods. [10,43,81] To introduce ion conducting functionality into polythiophene, we attach polar oligoethylene glycol side-chains to the polythiophene backbone. The originally proposed chemistry, P3MEET (left side

of Figure 2.1a), consists of a thiophene backbone and an oligoethylene glycol side-chain with an oxygen atom directly bonded to the thiophene ring. Although placing the electron donating oxygen atom right next to the thiophene core is known to reduce the ionization energy and improve electrochemical stability of the polymer,[103,113,143] it could negatively impact the ionic conductivity due to the reduced segmental motion of the oxygen

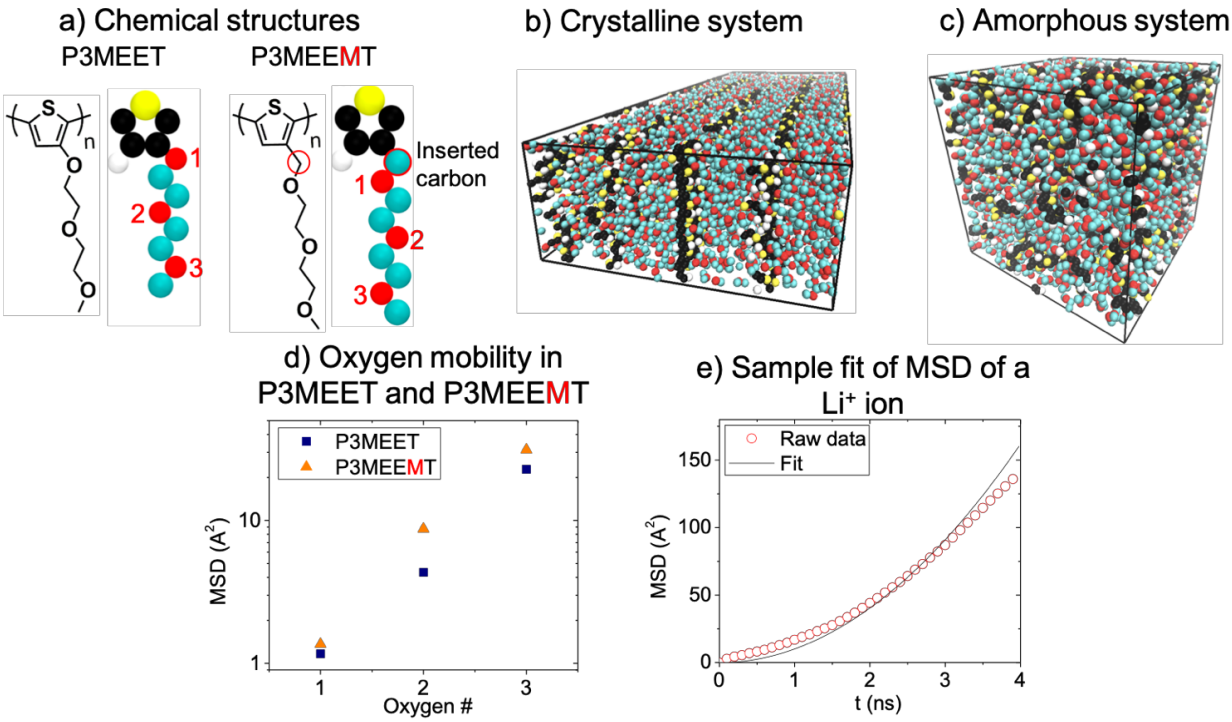

Figure 2.1: (a) Chemical structures of P3MEET and P3MEEMT. (b) (c) Snapshots of equilibrated crystalline and amorphous systems used in this study. (d) Long-time plateau mean square displacement (MSD) of 3 oxygen atoms in the side-chain of P3MEET and P3MEEMT. The oxygen # is noted in (a). (e) A typical plot of the MSD of a mobile Li^+ ion within P3MEEMT crystalline system and the fit used to calculate ion mobility.

atom covalently bonded to the rigid polythiophene backbone. Thus, to improve ion transport, a second chemistry, P3MEEMT (right side of Figure 2.1a), is proposed. P3MEEMT is chemically very similar to P3MEET but possesses a methylene spacer between the backbone and the first oxygen atom of the oligoethylene glycol chain.

Due to the semicrystalline nature of conjugated polymers, we use MD simulations to predict ion conductivity in the two limiting scenarios of purely crystalline and amorphous systems. To simulate the systems of interest, a hybrid all-atom (AA)/united-atom (UA) model is used where the ring is simulated with the OPLS-AA forcefield, while the side chains are modeled using the OPLS-UA forcefield, as shown in Figure 2.1a. The AA model is used for he thiophene rings as it is better suited to capture the ring-ring stacking. The side chains are modeled with the UA model for computational efficiency. Parameters for the model are given in the supplementary information (SI).

Different methods were used to obtain equilibrated structures of the phases of interest. For the crystalline phase, the initial configuration was equivalent to the crystalline structure of poly(3-hexylthiophene) (P3HT) (Isomorph I)[164] except that the crystals of both P3MEET and P3MEEMT were expanded to accommodate the longer side chains. The system was thermalized at 300 K and 1 atm (using the Nosé-Hoover barostat[143]) until metrics such as $E_{non\text{-}bonded}$, E_{bonded}, and the pressure tensor components become time-independent. For amorphous configurations, a CG model[16] which represents the ring as one bead, and the hexyl side chain as two beads was adapted to model poly-(3-nonyl-thiophene) (P3NT) by adding another bead to the side. Once the CG systems were equilibrated, an in-house reverse coarse graining code was used to generate a configuration of the UA/AA model. This configuration was then allowed to also equilibrate at 300K and 1atm. Snapshots of the equilibrated crystalline and amorphous morphology of P3MEEMT are given in Figure 2.1: (a) Chemical structures of P3MEET and P3MEEMT. (b) (c) Snapshots of equilibrated crystalline and amorphous systems used in this study. (d) Long-time plateau mean square

displacement (MSD) of 3 oxygen atoms in the side-chain of P3MEET and P3MEEMT. The oxygen # is noted in (a). (e) A typical plot of the MSD of a mobile Li+ ion within P3MEEMT crystalline system and the fit used to calculate ion mobility.b and c. Crystalline and amorphous structures of P3MEET appear similar and are thus not shown.

Former studies on the ionic conductivity in polymer electrolytes have correlated higher chain segment mobility with greater ionic conductivity.[21,227] Thus, after equilibrating the crystalline and amorphous structures, we investigated the segmental mobility of P3MEET and P3MEEMT side-chains by calculating the mean square displacement (MSD) of the three oxygen atoms in their side-chains as depicted in Figure 2.1a.We expectedly found that the MSD of each atom increases and then plateaus beyond a certain time period since the side-chain is covalently bound to the rigid polythiophene backbone (see Supporting Information). Figure 2.1d shows the plateau MSD of the three oxygens in ion-free P3MEET and P3MEEMT crystalline systems. For both polymers, the MSD of the oxygen atom expectedly increases with greater distance (from the anchoring thiophene backbone) along the oligoethylene side chain. Moreover, adding the carbon spacer in P3MEEMT increases the MSD, and hence the segmental mobility of oxygen atoms at corresponding positions relative to P3MEET for both crystalline and amorphous systems.

To calculate the ionic conductivity, free Li^+ ions were added to the systems at a concentration $r = [Li^+]/[EO] = 0.02$ and their mobility tracked upon the application of a large electric field of 500MV/m. The choice of r is such that ions are sufficiently dilute to minimize ion-ion interactions, but also abundant enough to allow for good statistical sampling. Free Li^+ ions are used instead of entire salts to avoid the need to simulate the very long time-scale associated with salt dissociation. This implicitly assumes a 100% dissociation ratio which is

a sensible assumption for LiTFSI in PEO-like environments. The electric field value chosen is large enough to allow for good sampling of the ionic motion (given the accessible simulation time scale), but still sufficiently moderate to yield a linear response behavior (Ohm's law). While applying such a large electric field, a tethering potential was used to keep the thiophene backbones near their original positions and avoid destroying the overall microstructure. Simulations were run for 30 ns in a constant volume ensemble with a set temperature of 400K to allow for better sampling. For calculating the ionic conductivity $\langle\sigma\rangle$, three replicate simulations were run and only "mobile" ions are considered, i.e., those showing a parabolic dependence of MSD on time as in the example plotted in Figure 2.1: (a) Chemical structures of P3MEET and P3MEEMT. (b) (c) Snapshots of equilibrated crystalline and amorphous systems used in this study. (d) Long-time plateau mean square displacement (MSD) of 3 oxygen atoms in the side-chain of P3MEET and P3MEEMT. The oxygen # is noted in (a). (e) A typical plot of the MSD of a mobile Li+ ion within P3MEEMT crystalline system and the fit used to calculate ion mobility.e. For such ions, the MSD dependence is fitted to the following equation from which the ionic mobility $<\mu>$ can be extracted:

$$MSD \approx (vt)^2 = (E\mu)^2 t^2 \qquad (2.1)$$

Where v, t, and E, are drift velocity, time, and electric field strength, respectively. The ionic conductivity, $\hat{\sigma}$, is then calculated as:

$$\hat{\sigma} = \langle\mu\rangle\langle N_{mobile}\rangle \quad (2.2)$$

where $\langle N_{mobile}\rangle$ is the average number of the "mobile" ions. More details of these conductivity calculations can be found in the Supporting Information. The average ionic conductivity of P3MEET and P3MEEMT in both crystalline and amorphous systems is summarized in Table 2.1.

Table 2.1: Calculated ion conductivity of P3MEG and P3MEET in equilibrated crystalline and amorphous systems from MD simulation

$\hat{\sigma}\left(cm^2/_{Vs}\right)$	Crystalline	Amorphous
P3MEEMT	5.0×10^{-6}	9.0×10^{-6}
P3MEET	2.4×10^{-5}	1.8×10^{-6}

Interestingly, while in the amorphous phase the conductivity is *ca.* 5 times higher in P3MEEMT than in P3MEET, the trend reverses in the crystal phase; i.e., the conductivity is ca. 5 times higher in P3MEET than in P3MEEMT. To explain this behavior, we show in Figure 2.2 a snapshot of P3MEEMT and P3MEET side-chain conformation in the presence of Li^+ ions within the crystalline domain. In pure PEO, two chains can form a "cage" around the ion, forming a deep kinetic trap for the ions.[21] A condition of cage formation is that the ion is only coordinated by two side chains, otherwise there are more chains coordinating the ion. To quantify cage formation, we determine the number side chains which have an atom within a 3.5Å cutoff of each ion and find the average number of atoms which are coordinating the ion. We find these values for P3MEET and P3MEEMT to be 2.47 ± 0.03 and 2.14 ± 0.003, respectively. This agrees with the idea that cage formation is higher in crystalline P3MEEMT, where the regular arrangement of side-chains and mobility of the oxygen near the backbone allow for cage formation. In contrast, in P3MEET the lack of mobility for the oxygen near the thiophene backbone reduces the chance of cage formation. In the amorphous systems, ion-trapping cages are more difficult to form because of the irregular arrangement of the side-

chains. Overall, our simulation results indicate the strong influence of side-chain chemistry on ionic transport characteristics of semicrystalline conjugated polymers.

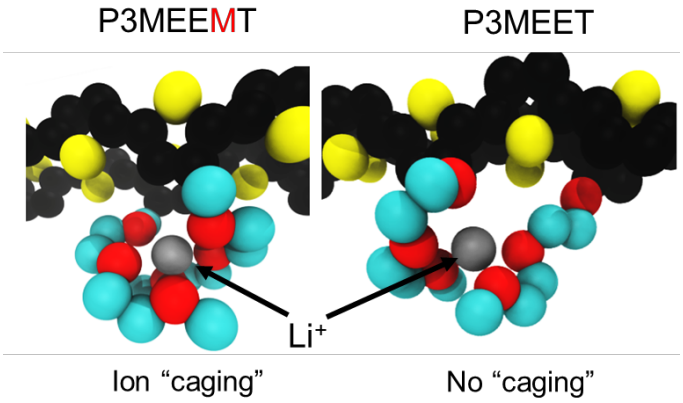

Figure 2.2: Snapshot showing the ion caging that is frequently seen in crystalline P3MEEMT but almost never in crystalline P3MEET.

2.3 Conclusion

In conclusion, we have performed a systematic study on the connection between morphological structure and ionic conduction characteristics of two polythiophene derivatives P3MEET and P3MEEMT upon blending with LiTFSI. MD simulations showed that oxygen atoms in P3MEEMT side-chains possessed higher segmental mobility compared to those of P3MEET, resulting in better Li^+ ion conductivity in amorphous phase. However, ionic conductivity of crystalline P3MEEMT was found to be lower than P3MEET due to the formation of Li^+ ion cages, a feature not observed in P3MEET. Using GIWAXS measurement, experimental collaborators showed that blending P3MEET and P3MEEMT with LiTFSI disturbed their crystalline structure, resulting in lattice strains along the side-chain stacking direction and significant drops in rDoC which stronger effect in P3MEET compared to

P3MEEMT. Introducing LiTFSI also exerted stronger influence on π-π interaction in P3MEET than in P3MEEMT, which most likely originated from the proximity of oxygen atom in P3MEET to the thiophene rings. Additionally, LiTFSI was found to infiltrate both the amorphous and crystalline domains at low concentrations but preferably resided in the amorphous domain at high LiTFSI concentrations. Temperature dependence studies of ionic conductivity at different concentration r in both P3MEET and P3MEEMT showed ionic conductivity increased from r = 0.01 to 0.05 but reduced at higher concentration due to ion pairing and transient cross-linking of PEO side-chains to Li^+ ions. Ionic transport was found to occur in amorphous regime, and always higher in P3MEEMT compared to P3MEET, consistent with our prediction using MD simulation.

Our results have demonstrated that side-chain modification can be a fruitful method to control ion transport in conjugated polymers. We also expect non-trivial effect of side-chain chemistry on electronic transport upon blending with ionic species, as LiTFSI has a different influence on rDoC and π-π interactions for P3MEET and P3MEEMT. Side-chain engineering is thus expected to be critical in the molecular design of mixed conduction materials. Finally, our approach using computational simulations to guide synthesis and characterization can be an elegant method towards studying the influence of chemistry on self-assembly, structure and ionic/electronic conduction of soft materials.

2.4 Acknowledgments

This work was supported by NSF DMREF Award Number 1629369. C.N. thanks Dr. Mayank Misra for help with DFT calculations.

2.5 Supplemental Information

2.5.1 Molecular Model

For atomistic simulations, we use the all-atom (AA) force field potential developed by Huang, *et. Al*[86] (based on the OPLS-AA forcefield[99]) to simulate the thiophene rings. Leveraging its similarity to the OPLS-AA model, we use the OPLS-UA forcefield[99] to model the side chains to decrease computational load. A depiction of our model is given in Figure 2.3. The non-bonded interactions are defined by a combination of Lennard-Jones (LJ) and Coulombic potential energy functions truncated at a cut-off of 10Å. The parameters for homonuclear non-bonded interactions are given in Table 2.2. We adopted the OPLS-AA forcefield prescription for the mixing rules for heteronuclear non-bonded interactions and scaling of non-bonded interaction for bonded atoms. Parameters for the bond and angle potentials are given in Table 2.3, and Table 2.4, respectively. For dihedral angle potentials, we used the standard OPLS potential adopted by the Cheung model with the parameters listed in Table 2.5. For the ring-ring and ring-side chain dihedrals we adopted the model and parameters of Poelking and Andrienko[164] given in Table 2.6. To account for the change in partial charges due to the change in chemistry, we altered the charges of the thiophene atoms and the first atom bonded to the ring to maintain charge neutrality and respect the relative values of Muliken charges calculated using BLYP 6-31G*. The remaining atoms in the side chains are assigned charges according to the OPLS-UA forcefield. Figure 2.3 displays all the atomic

charges whose values are assumed fixed (i.e., polarization effects are neglected based on the values of their dipole moments[21,223]).

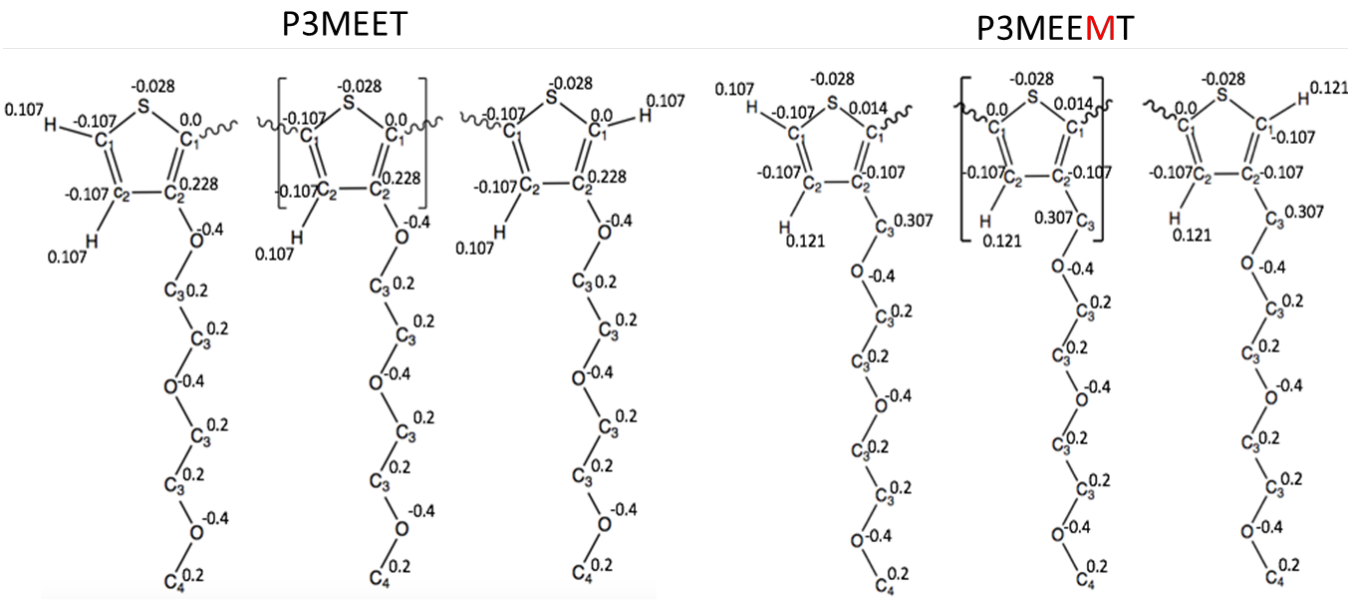

Figure 2.3: Depiction of the end, and inner monomers of P3MEET and P3MEEMT. Charges of the atoms are given.

Table 2.2: Non-boned parameters used in the AA model. The atom types are defined in Figure 2.3

Atom type	ε / (Kcal/mol)	σ / Å
C_1[a]	0.07	3.55
C_2[a]	0.07	3.55
C_3[b]	0.118	3.8
C_4[b]	0.17	3.8
H[a]	0.03	2.5
O[b]	0.17	3.0
S[a]	0.25	3.55

[a]Adapted from Ref. [86]
[b]Adapted from Ref [99]

Table 2.3: Harmonic bond parameters, where the potential is defined by harmonic potential; $U_{bond} = k_b(r - r_0)^2$

Type	k_b / (Kcal / (mol·Å^2))	r_0 / Å
C_1-C_2[a]	514.27	1.364
C_2-C_2[a]	453.1	1.428
C_1-C_1[a]	392.29	1.43
S-C_1[a]	291.25	1.732
C_1-H[a]	370.63	1.08
C_2-H[a]	370.63	1.08
C_2-C_3[b]	317.	1.51
C_2-O[b]	450.0	1.364
C_3-C_3[b]	260.0	1.526
C_3-O[b]	320.0	1.425
O-C_4[b]	320.0	1.425

[a]Adapted from Ref. [86]
[b]Adapted from Ref [99]

Table 2.4: Harmonic angle parameters, where the potential is defined by harmonic potential; $U_{angle} = k_\theta(\theta - \theta_0)^2$

Type	k_θ / (Kcal / (mol·radian^2))	θ_0 / degree
C_1-S-C_1	86.36	92.774
S-C_1-C_2	86.36	110.292
C_1-C_2-C_2	39.582	113.322

S-C_1-C_1	41.74	119.569
C_1-C_1-C_2	35.263	130.14
S-C_1-H	28.787	124.609
C_2-C_1-H	35.263	125.1
C_1-C_2-H	35.263	122.979
C_2-C_2-H	35.263	123.7
C_2-C_2-C_3	70.0	123.7
C_1-C_2-C_3	70.0	122.98
C_2-C_2-O	70.0	123.7
C_1-C_2-O	70.0	122.979
C_2-C_3-O	50.0	109.5
C_2-O-C_3	75.0	111.0
C_3-O-C_3	100.0	111.8
O-C_3-C_3	80.0	109.5
C_3-O-C_4	100.0	111.8

[a]Adapted from Ref. [86]
[b]Adapted from Ref [99]

Table 2.5: OPLS torsional parameters, where the potential is defined as; $U_{dihedral} = \sum_{n=1}^{4} \frac{1}{2} V_n[1 + (-1)^{n+1} \cos n\phi]$

Type	$V_1\left(\frac{1Kcal}{mol}\right)$	$V_2\left(\frac{Kcal}{mol}\right)$	$V_3\left(\frac{Kcal}{mol}\right)$	$V_4\left(\frac{Kcal}{mol}\right)$
C_1-S-C_1-C_2[a]	0	9.51	0	0
S-C_1-C_2-C_2[a]	0	9.51	0	0

C_1-C_2-C_2-C_1[a]	0	9.51	0	0
C_2-C_3-O-C_3[b]	0.65	-0.25	0.67	0
C_3-O-C_3-C_3[b]	0.65	-0.25	0.67	0
O-C_3-C_3-O[b]	-.55	0	0	0
C_3-C_3-O-C_4[b]	0.65	-.25	0.67	0
C_2-O-C_3-C_3[b]	0.65	-.25	0.67	0

[a]Adapted from Ref. [86]
[b]Adapted from Ref [99]

Table 2.6: Multi-Harmonic torsional parameters, where the potential is defined by harmonic potential; $U_{dihedral} = \sum_{n=1}^{6} V_n \cos^{n-1}\phi$

Type	$V_1\left(\frac{Kcal}{mol}\right)$	$V_2\left(\frac{Kcal}{mol}\right)$	$V_3\left(\frac{Kcal}{mol}\right)$	$V_4\left(\frac{Kcal}{mol}\right)$	$V_5\left(\frac{Kcal}{mol}\right)$	$V_6\left(\frac{Kcal}{mol}\right)$
S-C_1-C_1-S[c]	5.2132	-0.0502	-3.4139	0.1124	-1.5909	0.1268
C_1-C_2-C_3-O[c]	0.1866	-0.9069	0.0550	1.2440	-0.0359	0.2321
C_1-C_2-O-C_3[c]	0.1866	-0.9069	0.0550	1.2440	-0.0359	0.2321

[c]Adapted from Ref [164]

2.5.2 Calculation of Ionic mobility

The ion mobility from equilibrium MD simulations can be found from the diffusion constant based on the Einstein mobility relation, $\mu = Dq/kT$. Here, μ is the charge mobility, q is the charge, k is the Boltzmann constant, T is the temperature, and D is the diffusion constant calculated through the mean-square displacement (MSD) using:

$$D = \frac{1}{6} lim_{t\to\infty} \frac{\delta}{\delta t} \langle X_0^2(t) \rangle \ (2.3)$$

However, this relationship is not valid in the presence of an external electric field as it will induce a directional drift of the ions. Hence, in the presence of a uniform (and unidirectional) electric field the total displacement is given by

$$\langle X_E^2(t) - X_0^2(t) \rangle = (vt)^2 \ (2.4)$$

Where v is the drift velocity in the direction of the applied field. The ionic mobility can be calculated from:

$$\mu = \frac{vt}{E} \ (2.5)$$

where E is the applied electric field. We hence used the above equation as the basis for our mobility calculations. Since the calculated self-diffusion of Li^+ ions in the system with no applied field is insignificant ($\sim 10^{-8} cm^2/s$) compared to the calculated drift velocities, the expression for calculating the drift velocity then reduces to:

$$\langle X_E^2(t) \rangle = (vt)^2 \ (2.6)$$

which simplifies the calculation. Figure 2.4 shows how the MSD of each ion is fitted to Eq. 2.4. Ions showing a good fit to that expression are considered "mobile", i.e., moving in a way that can be described as under the influence of an electric field. For each mobile ion the fit is used to extract v and the average mobility is calculated from those values. Using MSD for calculating μ, and $\hat{\sigma}$ is based on convenience where the data for ion positions is easier to handle than using Green-Kubo relations. Previous studies which look at ionic conductivity use both methods.[21,223]

Ions that do not fit the form of Eq. 2.4 indicate that they have likely encountered a kinetic trap and hence are no longer experiencing continuous drift. In principle, the effect of those trapping events on the average ion mobility could be accounted for (and modelled) if

simulations could span long enough periods to collect suitable statistics of trapping-untrapping ("hopping") events. Since that is not the case, we leverage the fact that such events are rare to assume that a hopping mechanism has a small (slowing-down) effect in our estimated mobilities. Additionally, if a curve is fit to all ions and the value of conductivity is calculated, the variance between replicas is increased compared to the variance of conductivity values calculated when omitting ions which do not fit the curve well. Thus, we use the omission method to reduce variability in the data.

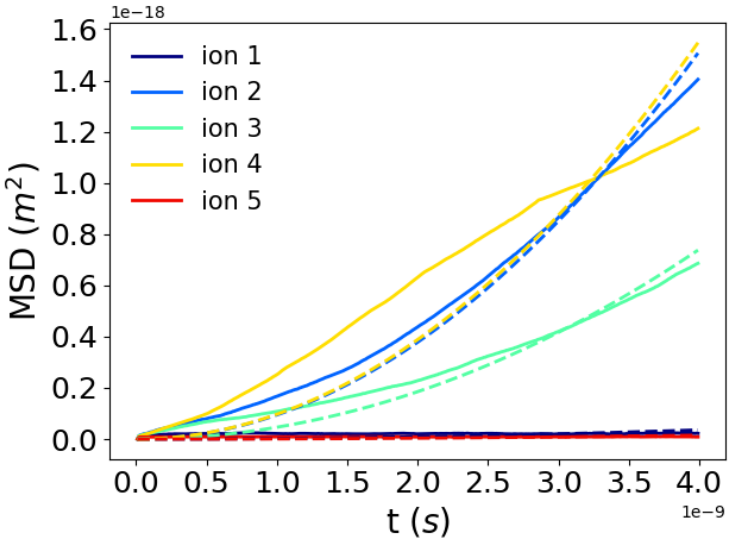

Figure 2.4: MSD plots of ions (solid lines) with the best fit to eq 2.4 for each ion plotted in the same color as a dashed line. Data is taken from the PMEEMT system in the crystalline phase.

2.5.3 Quantifying mobility of side chain oxygens

The data from Figure 2.1d is derived from the average plateau value of MSD for the each of the three oxygens in the side chains of P3MEET and P3MEEMT. Figures 2.5, 2.6, 2.7, and 2.8 show the average MSD vs. time plots for amorphous P3MEET, amorphous P3MEEMT,

crystalline P3MEET, and crystalline P3MEEMT, respectively. These plots are used to derive the plateau values (reached at long times).

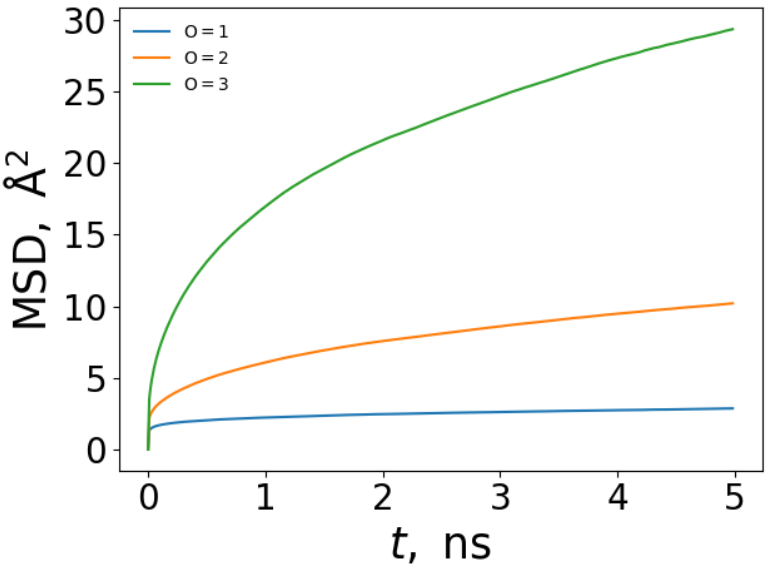

Figure 2.5: Average MSD plots for side chain oxygens in amorphous P3MEEMT (1 is closest to backbone, 3 is furthest from backbone).

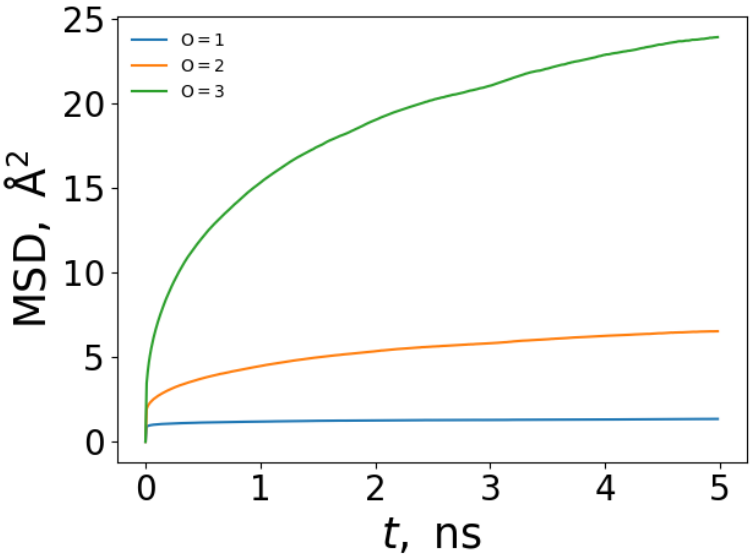

Figure 2.6: Average MSD plots for side chain oxygens in amorphous P3MEET (1 is closest to backbone, 3 is furthest from backbone).

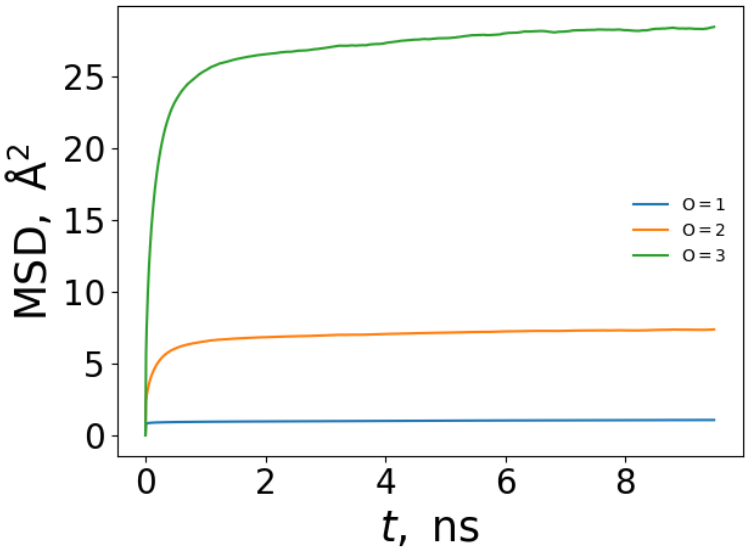

Figure 2.7: Average MSD plots for side chain oxygens in crystalline P3MEET (1 is closest to backbone, 3 is furthest from backbone).

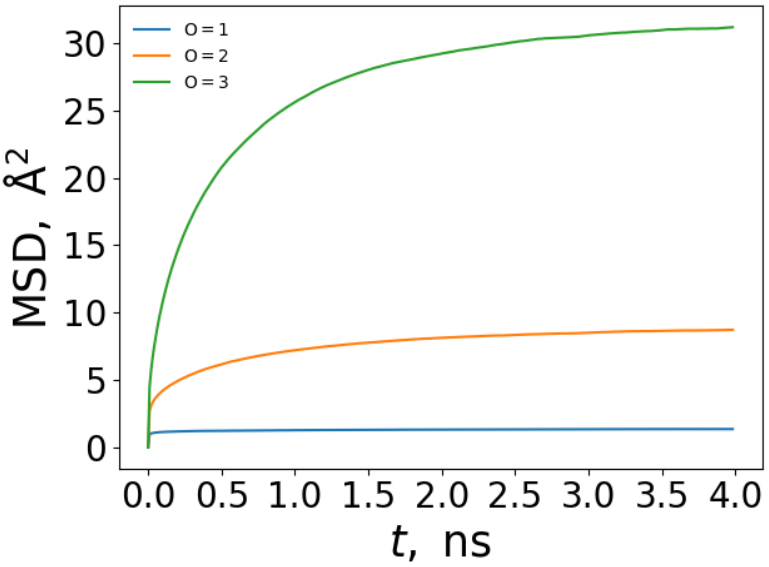

Figure 2.8: Average MSD plots for side chain oxygens in crystalline P3MEEMT (1 is closest to backbone, 3 is furthest from backbone).

3 Influence of Mesophases on Ionic Conduction Characteristics of Different Polythiophene Derivatives

3.1 Introduction

Conjugated polymers are of interest in applications such as xerography,[22] organic light-emitting diodes (OLEDs),[66,140] organic thin-film transistors (TFTs)[189] and organic photovoltaics (OPVs). [76,114] More recently, electrochemical applications have emerged requiring simultaneous electronic and ionic conducting capability raging from bioelectronics[104,123,154,173] to energy applications including batteries,[40,180] pseudo capacitors,[126] electrochromic windows,[139] dye-sensitized solar cells[167], and mechanical actuators.[77] The "soft" nature of conjugated polymers means that they will not damage any biological entity they are interfaced with, enabling a high quality interface, making them particularly attractive for bioelectronics applications.[90,173]

The key performance metrics of mixed ionic/electronic polymeric conductors are the electronic and ionic transport, which are strongly bound to the chemistry of the polymers

themselves. Ionic transport in conjugated polymers has been explored only on a limited basis. Most of the ion transport studies on conjugated polymers have been performed on poly(3,4-ethylenedioxythiophene) doped with poly (styrene sulfonate) (PEDOT: PSS) which has a rather complicated structure, preventing its use as a model system for structure-function relationship studies.[89,172,196] Despite our work[46] exploring how side chain chemistry influences the ionic conductivity behavior of polythiophene derivatives and recent work focusing on developing "non-PEDOT" conjugated polymers having polar side-chains and ion-conducting functionality,[68,143,162] detailed investigations on the influence of polymer chemistry on ion uptake, polymer morphology and ionic transport are still lacking. Therefore, designing high performance mixed conducting polymers requires an understanding of the relationship between polymer chemistry and the effect it has on ionic transport characteristics. We have shown that the simple insertion of a methylene group in the side chain chemistry of polythiophene derivatives can lead to order of magnitude changes in the ionic conductivity, highlighting the need for further studies.

To this end, we report an investigation of 6 new polythiophene derivates using Molecular Dynamics (MD). Based on poly(2,5- dibromo-3-(methoxyethoxyethoxymethyl)) (P3MEEMT), these 6 new designs span a range of oxygen concentration and arrangements to probe many different design principals. When the oxygen concentration in the side chain is reduced the energetic barrier to escape a solvation site is also reduced, however, factors such as percolation are negatively impacted, leading to a competition. Due to the ordered arrangement in the crystalline domain this competition is of lesser importance so design principles for maximizing ionic conductivity are found to differ between the crystalline and

amorphous phases. However, the amorphous domain is the primary phase where ionic conduction takes place, so we identify a side chain design which shows a ~5 times improvement in ionic conductivity in the amorphous phase compared to P3MEEMT. This provides a promising avenue for experimental collaborators to explore.

3.2 Model and Methods

The system preparation and ionic conductivity calculation methodology are identical to our past work (Chapter 2). The only difference between this work and past work is the chemistries tested which is captured in the Supplemental Information (SI).

3.3 Results & Discussion

Based on our past success with predicting the ionic conductivity of P3HT-like systems in Chapter 2 we seek to explore a broader range of side chain chemistries based on P3MEEMT. As Outlined in Figure 3.1 6 new chemistries are proposed, each one replacing some number

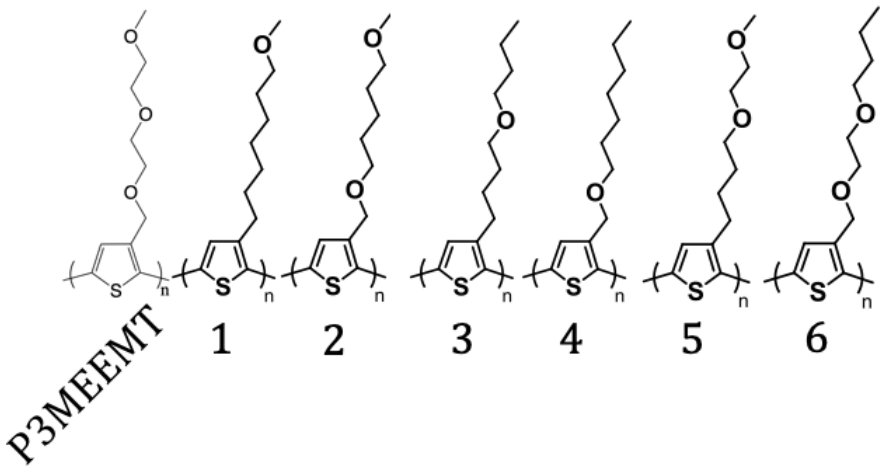

Figure 3.1: The chemistries evaluated in this study along with P3MEEMT, the previously identified material with good ionic conductivity, for comparison. The designs are based off of P3MEEMT with the only modifications being the replacement of certain oxygens in P3MEEMT with methylene groups.

of oxygens in the side chain of P3MEEMT with a methylene group. All possible configurations

(except for purely alkyl side chain) are considered to understand the effect of the number and positioning of charged groups in the side chain on ionic conductivity. We employ Molecular Dynamics simulations to study the ionic conductivity in purely crystalline and purely amorphous systems for each chemistry.

Ionic conductivity, σ, is calculated as:

$$\sigma = qvn \quad (3.1)$$

Where q, v, and n are the valence, drift velocity, and number density of dissociated ions, respectively. Our previous work compared two very similar chemistries where the number and relative positioning of oxygens was the same, so the concentration of dissociated ions was assumed equal for both systems. In contrast, the proposed side chains have differences in oxygen content and positioning so there will also be a difference in the concentration of dissociated salt present, which needs to be accounted for.

The dissociation process of a salt is controlled by the dielectric constant[62], ε, and the activation energy of the dissociation process, E_a, in an Arrhenius expression:

$$rate \propto e^{\frac{-E_a}{\varepsilon k_b T}} \quad (3.2)$$

Where k_b, and T are the Boltzmann constant and Temperature, respectively. A given salt pair (such as LiCl, adopted here for simplicity) will have a corresponding E_a and what changes is the value of ε for a given environment. One can take the ratio of E_a and ε to be a "dissociation energy", E_d, and propose a modified equation for σ:

$$\hat{\sigma} = \frac{qvn}{E_d} \quad (3.3)$$

A greater value of E_d means that there will be fewer dissociated ions, so E_d in the denominator of Eq. 3.3 mimics the effect of changing ε. While calculating ε directly for each

environment is feasible, it is a computationally demanding procedure. Here we calculate an effective E_d based on a potential of mean force, PMF. Hence, we seek to calculate the PMF (and E_d) of LiCl dissociation for each of the proposed polymer chemistries. To calculate the PMF, a series of simulations are conducted with the cation and anion (Li^+ and Cl^-) held at different specific separation distance, r_{sep}. For isotropic environments such as bulk water, the orientation of the salt pair when it is placed in the simulation box is unimportant. In contrast, the environment of the crystalline phase of polythiophene derivatives is anisotropic and the preferred orientation of the salt pair is unknown so it should be allowed to freely self-adjust. As such, to obtain the PMF, we employ the method outlined in Figure 3.2. After being placed in the system, Li^+, and Cl^- are modeled as a rigid body, i.e. the sum of the forces acting on Li^+, and Cl^-, are totaled and applied to the pair as if it were one rigid body while maintaining the distance between the two ions as implemented in previous work[101]. In this way r_{sep} remains constant while still allowing the salt pair to equilibrate to the preferred orientation. After the system is allowed to equilibrate, the mean force, MF can be calculated for the given r_{sep} as[60]:

$$MF(r_{sep}) = \langle(\vec{F}_{Li^+} - \vec{F}_{Cl^-}) \cdot \vec{r}\rangle \qquad (3.4)$$

Where $\vec{F}_i$, is the force vector acting on ion i, and $\vec{r}$ is the instantaneous unit vector between the ions. Simulations are run for r_{sep} from 2Å to 15Å, in increments of 0.2 Å, over 0.5ns to get an average value of PMF for the given r_{sep}. The energy required to take an ion pair from a given r_{sep} to an infinite separation, the PMF, can be calculated by[60]:

$$\mathrm{PMF}(r_{sep}) = \int_{r_{sep}}^{r_{max}} MF(r)\, dr \qquad (3.5)$$

Where r_{max} is the maximum value of r_{sep}. As seen in the bottom right of Figure 3.2 the value of r_{max} is chosen to be large enough to see a plateau in the PMF. Finally, all values are shifted so PMF(r_{max})=0 for clarity. The value of E_d is taken as the difference between the largest and smallest values of the PMF.

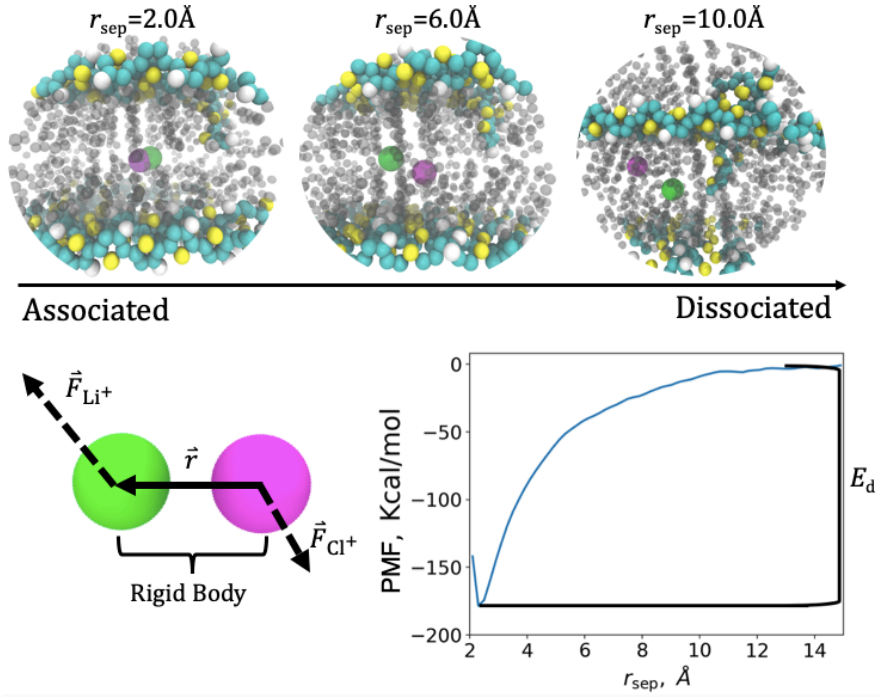

Figure 3.2: top: A schematic of how the distance between Li^+ and Cl^- is increased, ranging from an associated state to a dissociated state. bottom: a schematic of the variables used in equation 3.4 to calculate the PMF, and an example plot of PMF integrated over r_{sep} to give the change in energy, ΔE, for an ion pair going from the dissociated state to a given r_{sep}. Carbon, Sulfur, Hydrogen, Lithium, and Chlorine are colored as cyan, yellow, white, green, and purple. The side chains are colored transparent grey for clarity.

Summarized in Figure 3.3 is the "Normalized Conductivity Score" where $\hat{\sigma}$ for a given chemistry in either mesophase is normalized to $\hat{\sigma}$ for P3MEEMT in the corresponding mesophase. The trends seen in the crystalline phase are controlled by the concentration and location of oxygen in the side chains. Oxygen concentration controls the energetic

barrier needed to jump from one solvation site to another. With greater oxygen concentration an ion will on average be solvated by more oxygen atoms, therefore increasing the ionic mobility. This would lead one to conclude that reducing the oxygen concentration is purely beneficial, however, this reduction in oxygen content also increases E_d (see SI) which lowers the amount of dissociated salt, so a competition arises. This competition is highlighted in the best performers for crystalline conductivity, designs 1 and 2, where design 1 has lower oxygen content than design 2, however both systems have similar conductivity because of the difference in E_d. Despite all designs having higher values of E_d because of the reduction in oxygen concentration compared to P3MEEMT, the conductivity for all designs (besides design 6) show improved conductivity. This is because the removal of just one oxygen in the side chain disrupts the ability of the side chains to "cage" the ion as was seen for P3MEEMT in our previous work. The exception to the universally improved conductivity, design 6, has very low conductivity because the ions are drawn closer to the thiophene rings where there is reduced mobility, leading to reduced conductivity even though the oxygen concentration is reduced.

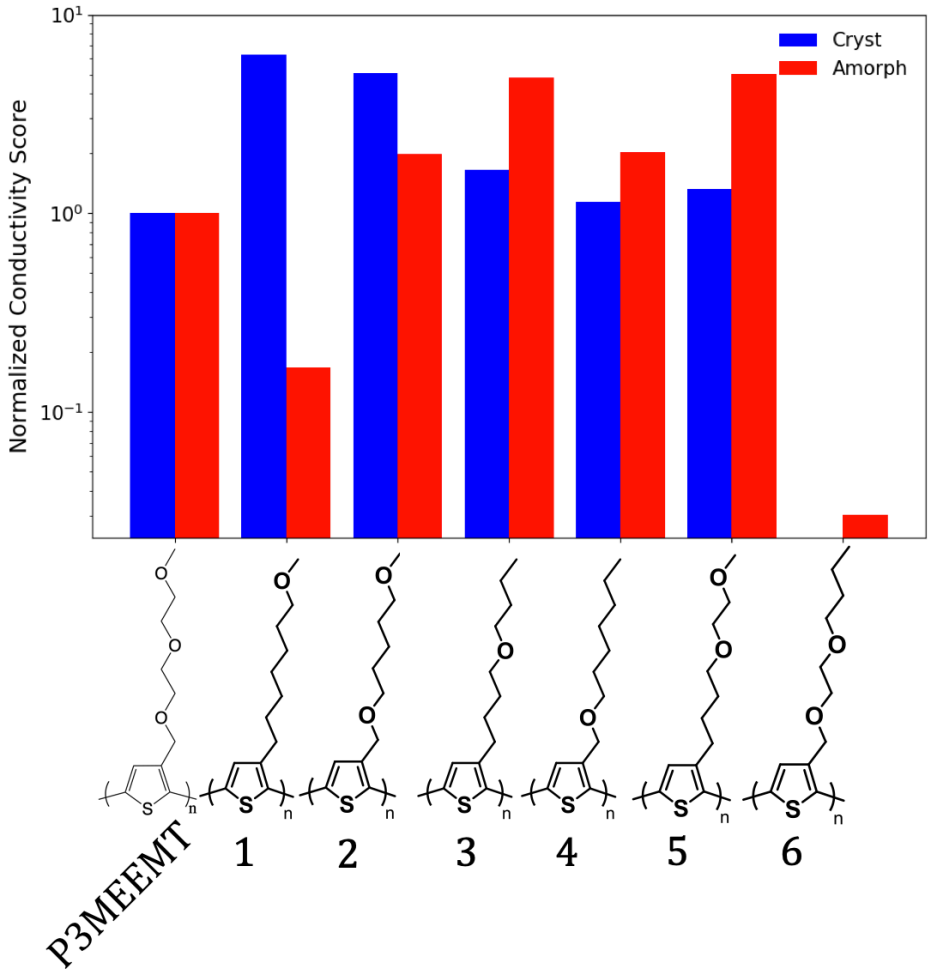

Figure 3.3: The conductivity scores for all proposed chemistries normalized to P3MEEMT for the respective mesophases.

The location of oxygens in the side chain is also vital, as the Li^+ ion is drawn to oxygens and those that are closer to the end of the side chain have greater mobility, as show in our previous work (Chapter 2). This can be seen in the difference in conductivities between designs 1, 3 and 4. These designs have identical oxygen content, and similar E_d, however, the greater mobility of the oxygen in design 1 leads to the greatest conductivity of the three designs, while design 4 has the worst conductivity because the oxygen closest to the thiophene ring has the lowest mobility.

Due to the regular arrangement of side chains in the crystalline phase, there is always a percolating pathway of solvation sites, so there is no need to consider percolation there. However, in the amorphous phase the loss of a regular arrangement necessitates the consideration of percolation. A reduction in oxygen content has the effect of increasing the average distance between oxygens, which increases the distance the ion needs to jump, leading to a greater energetic barrier, and therefore, lower conductivity[160,161]. To evaluate how the designs affect the distribution of oxygens in the amorphous phase we use a clustering algorithm which groups oxygens together as part of the same cluster if they are within a given cutoff, r_{cut}. The value of r_{cut} for which each system shows only one cluster, r_{cut}^{*}, can be thought of as the distance the ion needs to jump to go from one solvation site to the next. The rate of an ion making a jump of size l can be modeled using an Arhennius expression so the drift velocity, v_{d}, of an ion can be approximated as:

$$v_d \propto le^{-l} \qquad (3.6)$$

Which is maximized at small l, and therefore , r_{cut}^{*}. The value of r_{cut}^{*} for each system is plotted in Figure 3.4. As expected, P3MEEMT with the greatest concentration of oxygen in the side chain has the smallest value of r_{cut}^{*}, and therefore the smallest distance the Li^{+} ion needs to jump between solvation sites. Designs 2, 5 and 6 all have the same oxygen concentration, and therefore have similar r_{cut}^{*}. A similar trend is seen for designs 1, 3 and 4. This difference in r_{cut}^{*} explains why designs 1 and 2, which have similar performance in the crystalline phase, have very different performances in the amorphous phase (almost an order of magnitude difference).

Our previous work indicated that the amorphous phase is where ionic conductivity mainly takes place. As such, Design 5 emerges as the best performer, as it strikes a balance between reducing oxygen content, drawing Li^+ to the side chain end, and maintaining good percolation of solvation sites. Another promising candidate is design 2, which shows a ~2 times increase in conductivity in the amorphous phase. While not as large of an improvement in the ionic conductivity, the oxygen which is close to the thiophene ring may bring electronic dopants closer to the ring, improving electronic conductivity.

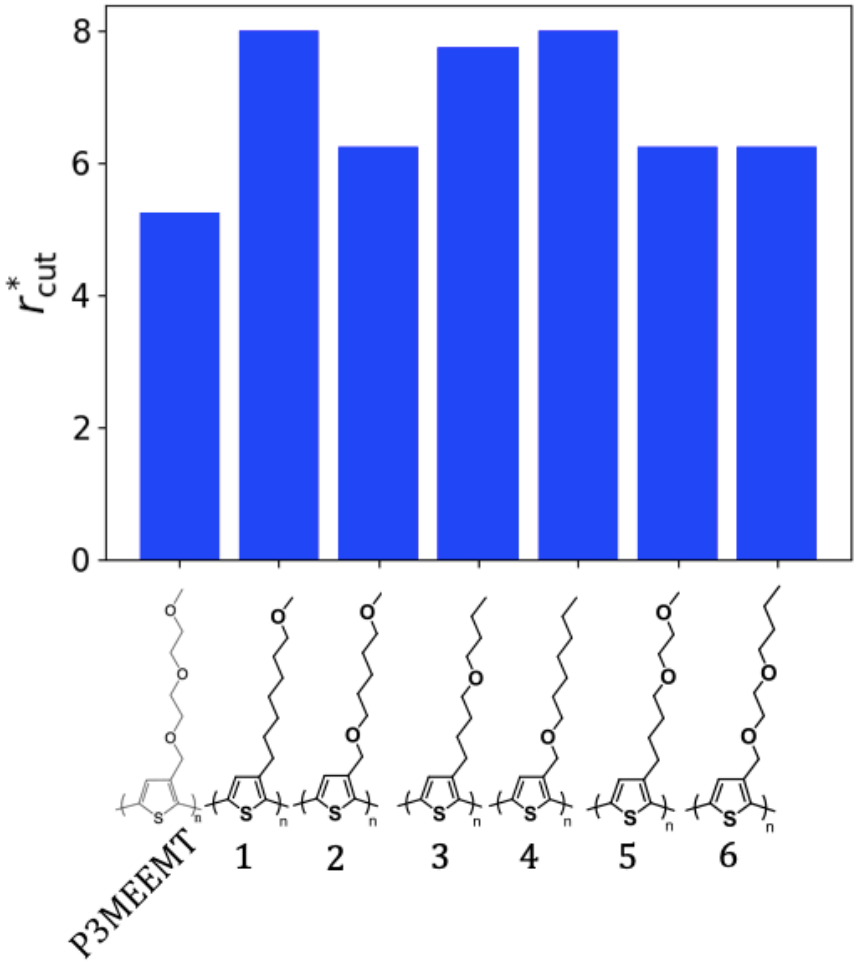

Figure 3.4: Values of r^*_{cut} for P3MEEMT and all proposed chemistries.

3.4 Conclusions

In conclusion, we have evaluated the performance of several different polythiophene derivatives based on P3MEEMT and identified promising candidates to investigate for improving upon the ionic conductivity of P3MEEMT. To account for the changing oxygen content in the side chains of the proposed chemistries, the energy of dissociating a given salt pair in the system of interest, E_d, is introduced into the calculation of conductivity. E_d can be obtained from the PMF which is calculated by modeling a salt pair as a rigid body, calculating the average mean force acting on the ions for different separation distances between the ions, and finally integrating these values. Modeling the salt pair as a rigid body allows for proper equilibration to an a priori unknown configuration.

Comparing the normalized conductivity scores of the new designs reveals that in the crystalline phase the designs which draw the ion to the end of the side chain and reduce the oxygen concentration are the best performers. In contrast the amorphous phase shows an opposing trend where reducing the oxygen concentration in the side chain, while reducing the energetic barrier to escape the solvation site, increases the average distance between solvation sites which leads to a reduced drift velocity, and therefore, conductivity, as predicted by a simple Arrhenius-like expression. The best performers in the amorphous phase balance the reduction in oxygen concentration while maintaining a short distance between solvation sites.

The conductivity in the amorphous phase is of primary importance as the amorphous phase is where ionic conduction primarily occurs. This leads to design 5 having the best predicted

ionic conductivity (~5 times improved amorphous conductivity over P3MEEMT). While not as good of a performer, design 2 (~2 times improved amorphous conductivity over P3MEEMT) has interesting implications when considering the dual conducting nature of these materials. Based on our results, experimental collaborators will seek to synthesize side chains of interest, namely design 5.

Future work will look to expand to a broader design space and employ machine learning approaches to help identify new high performing candidates. To do this, we plan to use a "library" of chemical functional groups which serve as "blocks" which are linked together to form new chain designs. The chemical identity and location of each block in the side chain design provide an easy way to represent a chain design to machine learning algorithms allowing for implementation of pre-existing algorithms such as the Genetic Algorithm, and Neural Networks. A Genetic Algorithm can guide incremental improvements in conductivity over several generations while A Neural Network can take in a large data set of the conductivity score for many side chain designs and make predictions about the entire design space. Both offer different advantages, as the Genetic Algorithm is easier to implement while the Neural Network can help predict far more chemistries than the Genetic Algorithm.

We will also seek to apply the method for calculating E_d to more complex salts such as LiTFSI, the primary salt used by experimental collaborators, to see if this change in salt affects the relative values of E_d as compared to the values obtained for LiCl. The anionic atom of interest in FSI^-, N, can modeled as a rigid body with Li^+ while the rest of the FSI^- ion can be allowed to freely equilibrate. This allows for a constant r_{sep} while also allowing for the salt to equilibrate

not only orientationally, but also configurationally, which is not a concern in simple diatomic salts like LiCl. Additionally, polarizable forcefields will be considered when calculating E_d for more complex salts since the polarizability of ions can greatly affect the PMF as the charge can be smeared across a complex ion such as FSI^- which isn't as prevalent for simpler ions such as Cl^-.

3.5 ACKNOWLEDGEMENTS

This work was supported by NSF DMREF Award Number 1629369. C.N. thanks Dr. Mayank Misra for insightful discussions

3.6 Supplemental Information

3.6.1 Simulation Model

For atomistic simulations, we use the all-atom (AA) force field potential developed by Cheung, *et. Al*[86] (based on the OPLS-AA forcefield[99]) to simulate the thiophene rings. Leveraging its similarity to the OPLS-AA model, we use the OPLS-UA forcefield[99] to model the side chains to decrease computational load. A depiction of our model is given in Figure 3.5. The non-bonded interactions are defined by a combination of Lennard-Jones (LJ) and Coulombic potential energy functions truncated at a cut-off of 10Å. The parameters for homonuclear non-bonded interactions are given in Table 3.1. We adopted the OPLS-AA forcefield prescription for the mixing rules for heteronuclear non-bonded interactions and scaling of non-bonded interaction for bonded atoms. Parameters for the bond and angle potentials are given in Table 3.2, and Table 3.3, respectively. For dihedral angle potentials, we used the standard OPLS potential adopted by the Cheung model with the parameters

listed in Table 3.4. For the ring-ring and ring-side chain dihedrals we adopted the model and parameters of Poelking and Andrienko[164] given in Table 3.5. To account for the change in

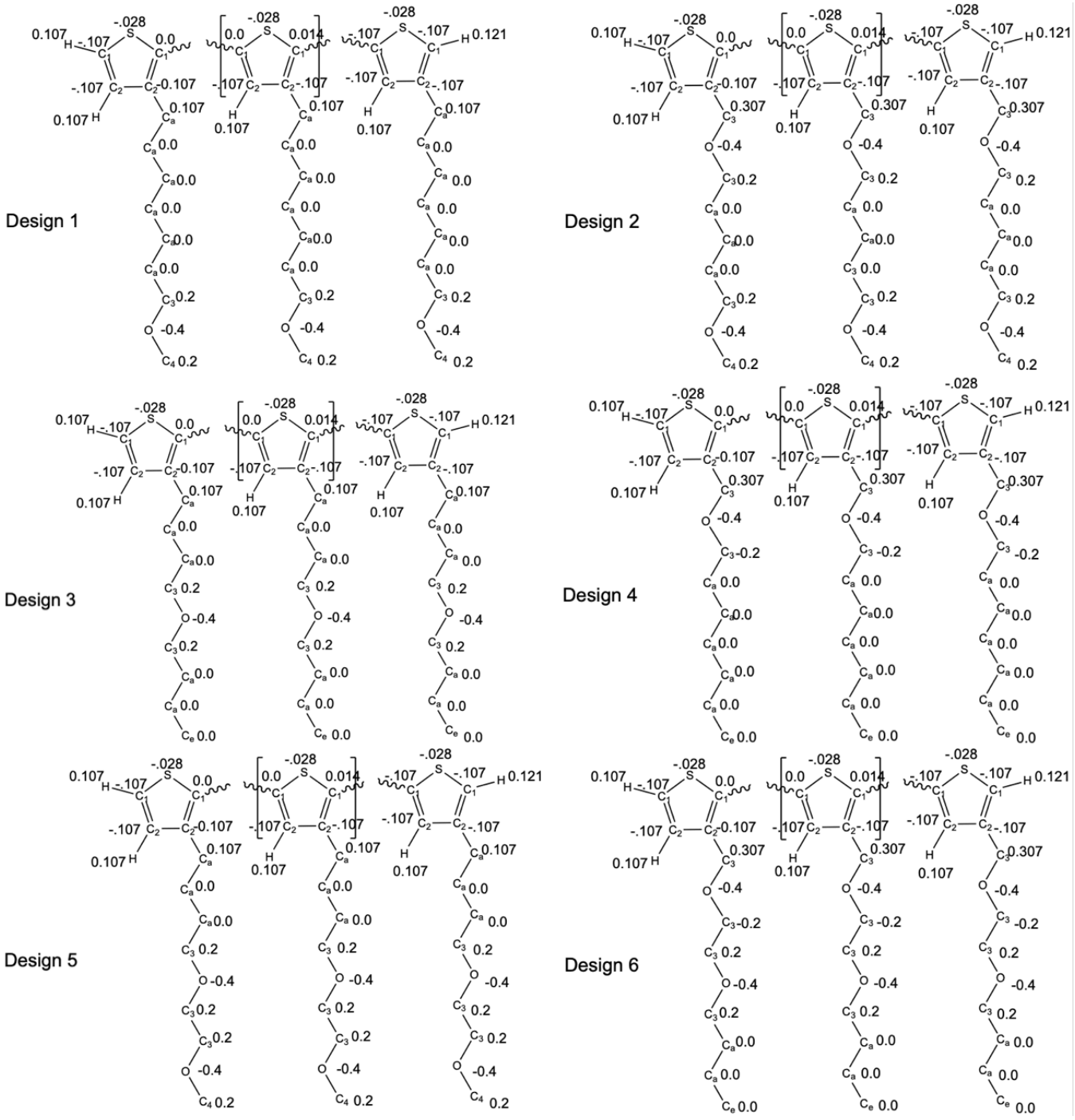

Figure 3.5: Depiction of the end, and inner monomers of P3MEET and P3MEEMT. Charges of the atoms are given.

partial charges due to the change in chemistry, we altered the charges of the thiophene atoms

and the first atom bonded to the ring to maintain charge neutrality and respect the relative values of Muliken charges calculated using BLYP 6-31G*. The remaining atoms in the side chains are assigned charges according to the OPLS-UA forcefield. Figure 3.5 displays all the atomic charges whose values are assumed fixed (i.e., polarization effects are neglected based on the values of their dipole moments[21,2223]).

Table 3.1: Non-boned parameters used in the AA model. The atom types are defined in Figure 3.5

Atom type	ε / (Kcal/mol)	σ / Å
C_1[a]	0.07	3.55
C_2[a]	0.07	3.55
C_3[b]	0.118	3.8
C_4[b]	0.17	3.8
C_a[b]	0.118	3.905
C_e[b]	0.175	3.905
H[a]	0.03	2.5
O[b]	0.17	3.0
S[a]	0.25	3.55

[a]Adapted from Ref. [86]
[b]Adapted from Ref [99]

Table 3.2: Harmonic bond parameters, where the potential is defined by harmonic potential; $U_{bond} = k_b(r - r_0)^2$

Type	k_b / (Kcal / (mol·Å^2))	r_0 / Å
C_1-C_2[a]	514.27	1.364

C_2-C_2[a]	453.1	1.428
C_1-C_1[a]	392.29	1.43
S-C_1[a]	291.25	1.732
C_1-H[a]	370.63	1.08
C_2-H[a]	370.63	1.08
C_2-C_3[b]	317.0	1.51
C_2-O[b]	450.0	1.364
C_3-C_3[b]	260.0	1.526
C_3-O[b]	320.0	1.425
O-C_4[b]	320.0	1.425
C_2-C_a[b]	317.0	1.51
C_a-C_ab	260.0	1.526
C_a-C_e[b]	320.0	1.425
C_a-C_3[b]	260.0	1.526

[a]Adapted from Ref. [86]
[b]Adapted from Ref [99]

Table 3.3: Harmonic angle parameters, where the potential is defined by harmonic potential; $U_{angle} = k_\theta(\theta - \theta_0)^2$

Type	k_θ / (Kcal / (mol·radian^2))	θ_0 / degree
C_1-S-C_1[a]	86.36	92.774
S-C_1-C_2[a]	86.36	110.292
C_1-C_2-C_2[a]	39.582	113.322
S-C_1-C_1[a]	41.74	119.569

C_1-C_1-C_2[a]	35.263	130.14
S-C_1-H[a]	28.787	124.609
C_2-C_1-H[a]	35.263	125.1
C_1-C_2-H[a]	35.263	122.979
C_2-C_2-H[a]	35.263	123.7
C_2-C_2-C_3[b]	70.0	123.7
C_1-C_2-C_3[b]	70.0	122.98
C_2-C_2-O[b]	70.0	123.7
C_1-C_2-O[b]	70.0	122.979
C_2-C_3-O[b]	50.0	109.5
C_2-O-C_3[b]	75.0	111.0
C_3-O-C_3[b]	100.0	111.8
O-C_3-C_3[b]	80.0	109.5
C_3-O-C_4[b]	100.0	111.8
C_2-C_2-C_a[b]	70.0	123.7
C_1-C_2-C_a[b]	70.0	122.979
C_2-C_a-C_a[b]	63.0	114.0
C_a-C_a-C_a[b]	63.0	112.4
C_a-C_a-C_3[b]	63.0	112.4
C_a-C_3-O[b]	63.0	112.4
C_a-C_a-C_e[b]	100.0	111.8

[a]Adapted from Ref. [86]
[b]Adapted from Ref [99]

Table 3.4: OPLS torsional parameters, where the potential is defined as; $U_{dihedral} = \sum_{n=1}^{4} \frac{1}{2} V_n[1 + (-1)^{n+1} \cos n\phi]$

Type	$V_1\left(\frac{1Kcal}{mol}\right)$	$V_2\left(\frac{Kcal}{mol}\right)$	$V_3\left(\frac{Kcal}{mol}\right)$	$V_4\left(\frac{Kcal}{mol}\right)$
C_1-S-C_1-C_2[a]	0	9.51	0	0
S-C_1-C_2-C_2[a]	0	9.51	0	0
C_1-C_2-C_2-C_1[a]	0	9.51	0	0
C_2-C_3-O-C_3[b]	0.65	-0.25	0.67	0
C_3-O-C_3-C_3[b]	0.65	-0.25	0.67	0
O-C_3-C_3-O[b]	-.55	0	0	0
C_3-C_3-O-C_4[b]	0.65	-.25	0.67	0
C_2-O-C_3-C_3[b]	0.65	-.25	0.67	0
C_a-C_a-C_a-C_3[b]	-3.4	1.25	3.1	0
C_a-C_a-C_3-O[b]	-.55	0	0	0
C_a-C_3-O-C_3[b]	0.65	-.25	0.67	0
C_a-C_3-O-C_4[b]	0.65	-.25	0.67	0
C_a-C_a-C_a-C_a[b]	-3.4	1.25	3.1	0
C_a-C_a-C_a-C_e[b]	-3.4	1.25	3.1	0
C_3-C_a-C_a-C_e[b]	-3.4	1.25	3.1	0

[a]Adapted from Ref. [86]
[b]Adapted from Ref [99]

Table 3.5: Multi-Harmonic torsional parameters, where the potential is defined by harmonic potential; $U_{dihedral} = \sum_{n=1}^{6} V_n \cos^{n-1}\phi$

Type	$V_1\left(\frac{Kcal}{mol}\right)$	$V_2\left(\frac{Kcal}{mol}\right)$	$V_3\left(\frac{Kcal}{mol}\right)$	$V_4\left(\frac{Kcal}{mol}\right)$	$V_5\left(\frac{Kcal}{mol}\right)$	$V_6\left(\frac{Kcal}{mol}\right)$

S-C_1-C_1-S[c]	5.2132	-0.0502	-3.4139	0.1124	-1.5909	0.1268
C_1-C_2-C_3-O[c]	0.1866	-0.9069	0.0550	1.2440	-0.0359	0.2321
C_1-C_2-C_a-C_a[c]	0.1866	-0.9069	0.0550	1.2440	-0.0359	0.2321

[c]Adapted from Ref [164]

3.6.2 E_d for all chemistries

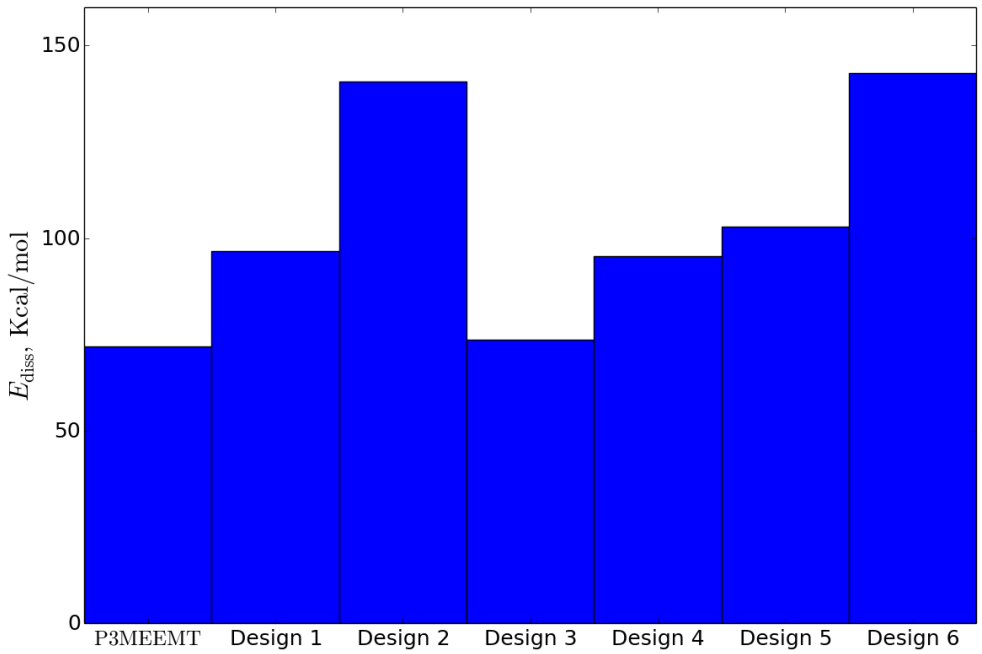

Figure 3.6: The values of E_d for all chemistries.

3.6.3 Atoms coordinating Li^+ ions

Atoms in the side chain are assigned numbers starting at 1 for the atoms bonded closest to the thiophene backbone, up to 9 for the last atom in the side chain. During the ionic conduction simulations, an atom in the side chain is considered to be coordinating the Li^+ ion if it is within 3.5Å. The probability histograms are plotted to see how each chain design coordinates the ion.

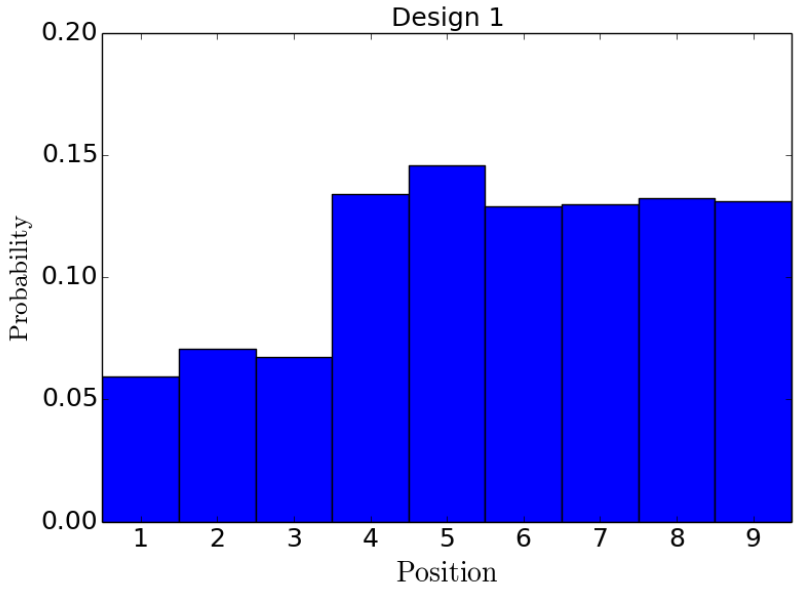

Figure 3.7: The probability histogram for atoms in the side chain of design 1 to be coordinating Li^+.

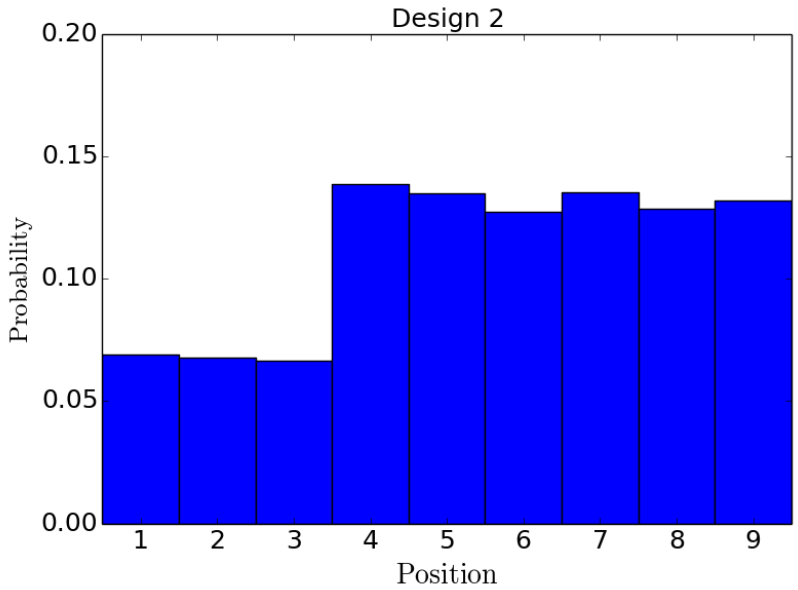

Figure 3.8: The probability histogram for atoms in the side chain of design 2 to be coordinating Li^+.

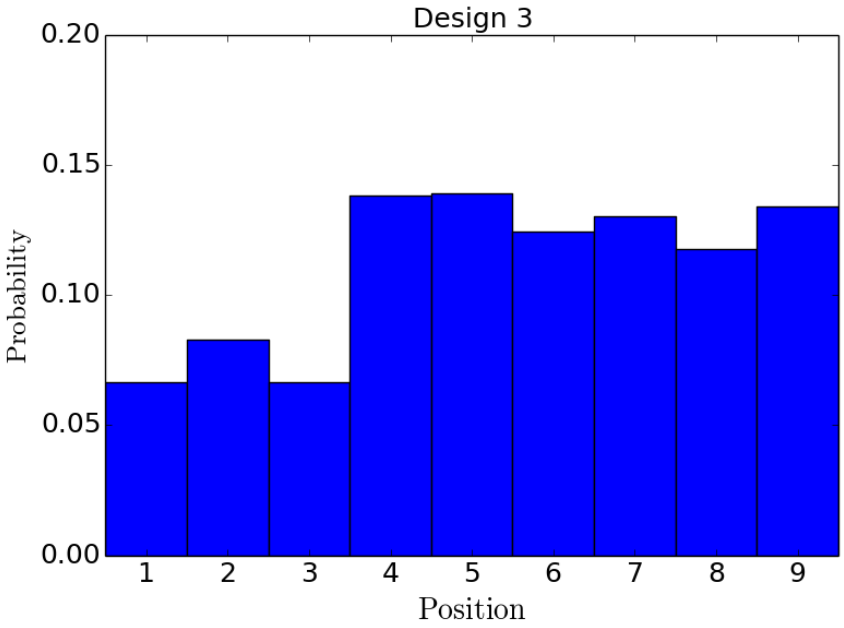

Figure 3.9: The probability histogram for atoms in the side chain of design 3 to be coordinating Li^+.

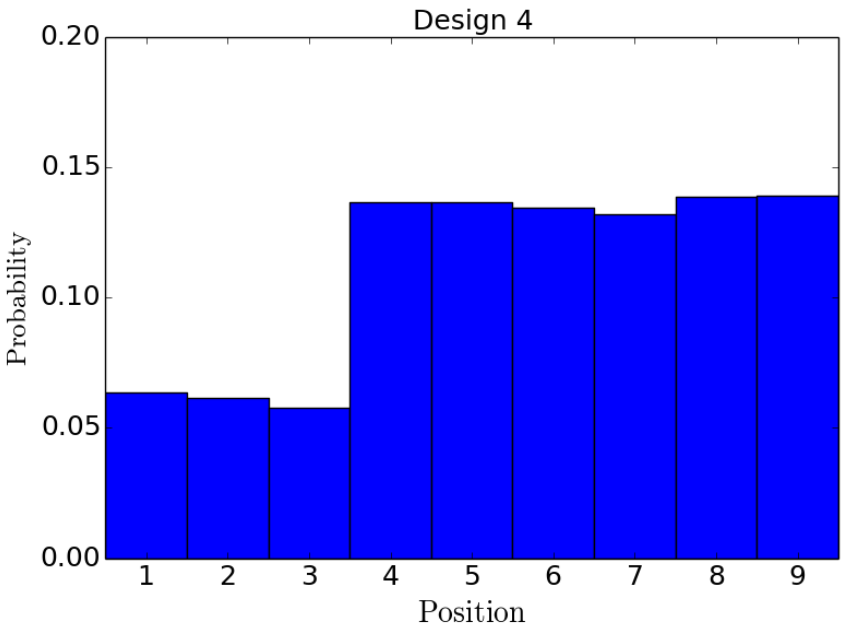

Figure 3.10: The probability histogram for atoms in the side chain of design 4 to be coordinating Li^+.

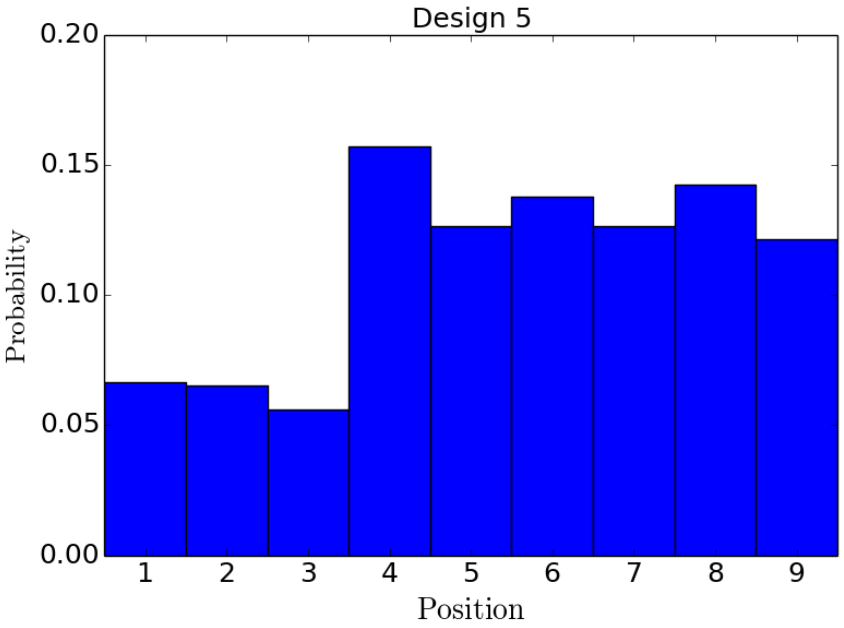

Figure 3.11: The probability histogram for atoms in the side chain of design 5 to be coordinating Li^+.

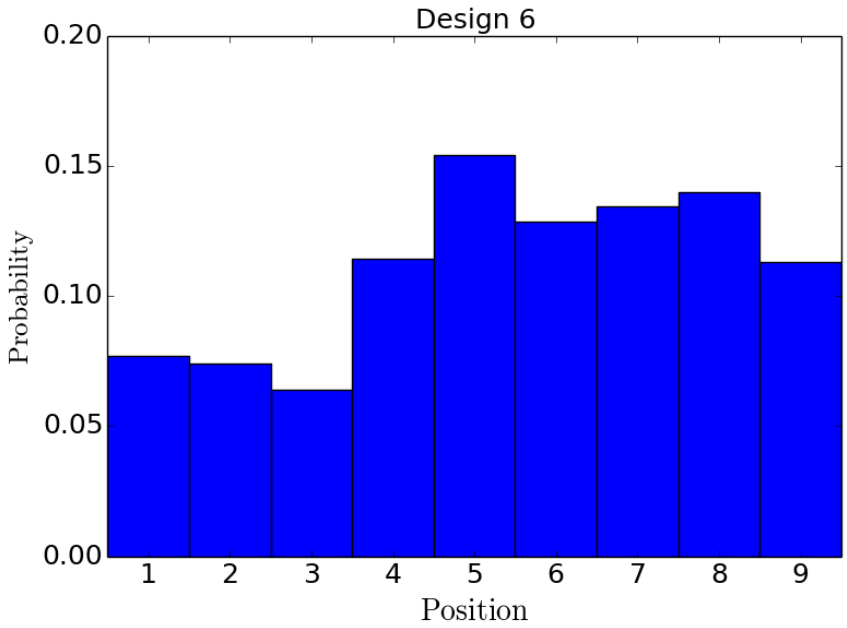

Figure 3.12: The probability histogram for atoms in the side chain of design 6 to be coordinating Li^+.

4 Optimizing the network topology of block copolymer liquid crystal elastomers for enhanced extensibility and toughness

4.1 INTRODUCTION

Elastomers, crosslinked networks of polymer chains, often possess high extensibility but low toughness, γ, the energy absorbed upon deformation before breakage. While increasing crosslinking density increases the young's modulus, it also decreases extensibility and may not necessarily enhance γ[11,192]. It is thus desirable but challenging to develop elastomers with high γ that are also highly extensible; such materials would be appealing for applications requiring super-toughness such as specialty fibers, shock absorbers, artificial muscles, etc. This combination of properties has been predicted to be exhibited by model liquid crystalline elastomers (LCE) with unentangled, tetra-functional cross-linking (T-LCE) and ABA block copolymers connecting the crosslinks [150,4]. These T-LCEs achieve that feat by producing a saw-tooth shaped stress-strain response to uniaxial deformation, wherein each "tooth" is associated with microscopic rearrangements that increase the number of AB layers along the strain axis. A key question to be addressed by this study is whether realistic crosslinking schemes that generate complex and defective connectivities[184], can also lead to elastomers with high toughness and extensibility.

LCE are elastomers where the constituent chains are semi-flexible, i.e., have a large persistence length, l_p~100Å, that form orientationally aligned liquid crystal phases induced by various stimuli[34,44,124,188,204]. In LCEs, network deformation is coupled to chain orientation[204], a feature that can be exploited in artificial muscles[125,2222] and high-strain actuators[218,232]. T-LCE are LCE where, under full 3D extension, tetra-functional

cross-links would sit on the lattice points of a diamond lattice with the homopolymers acting as the carbon bonds in diamond. Due to main-chain backbone semi-flexibility, the chains and cross-links self-assemble into separate layers (Figure 4.1) [4,5,17]. Upon uniaxial deformation, stress builds in these networks until chain backbone hairpins unfold and a new cross-link layer forms. For semi-flexible polymers, the mechanism of chain unfolding entails "shifting" the mass of one hairpin leg to the other leg until the first leg is short enough that the bending energy overcomes entropy and fully unfolds[226]. Upon hairpin unfolding, the network "yields", to reduce chain overstretching, and a new layer is formed. For T-LCE, hairpin unfolding is further hindered by the cross-linking of chain ends, and the entropic segregation of the cross-links that is broken during unfolding. For T-LCE with ABA block copolymer chains, the addition of an interfacial energy penalty enhances the segregation, and favors the formation of smectic (orientationally and translationally ordered) domains of A and B blocks [4,150]. While forming a new domain, the surface area, and consequently the interfacial energy, increases, requiring larger stresses for domain formation as compared to homopolymer T-LCE. Further deformation leads to a succession of large stress peaks giving the characteristic saw-tooth stress-strain pattern [4,150]. This deformation mechanism is reminiscent of naturally occurring tough materials such as titin, and spider silk, both of which undergo non-affine domain deformations to relieve stress[18,91,109,148,171,230]. The concept of yielding to allow for greater energy absorption is the cornerstone of the "crumple zone", a concept used in cars and protective equipment[79,100].

T-LCEs embody a particular type of model network topology wherein tetrafunctional sites are interconnected as in a single diamond lattice, but the saw-tooth tensile response is not expected to be a unique characteristic of such a topology. In fact, other LCE topologies with

regular lattice connectivity such as double gyroid and double diamond have also been shown to produce saw-tooth tensile responses [5]. But such model LCEs (including T-LCE) represent the limiting case of perfect, unentangled networks; in contrast, both natural and synthetic LCEs exhibit topological inhomogeneities and entanglements that affect the tensile response[18,19,28,38,211]. Indeed, entanglements tend to impede cross-link mobility which is vital to the formation of multiple domains, and reduce network extensibility α_m= maximum strain before stress divergence (when further domain formation does not restore the stress to ~ 0), limiting the amount of energy that can be absorbed during deformation [210].

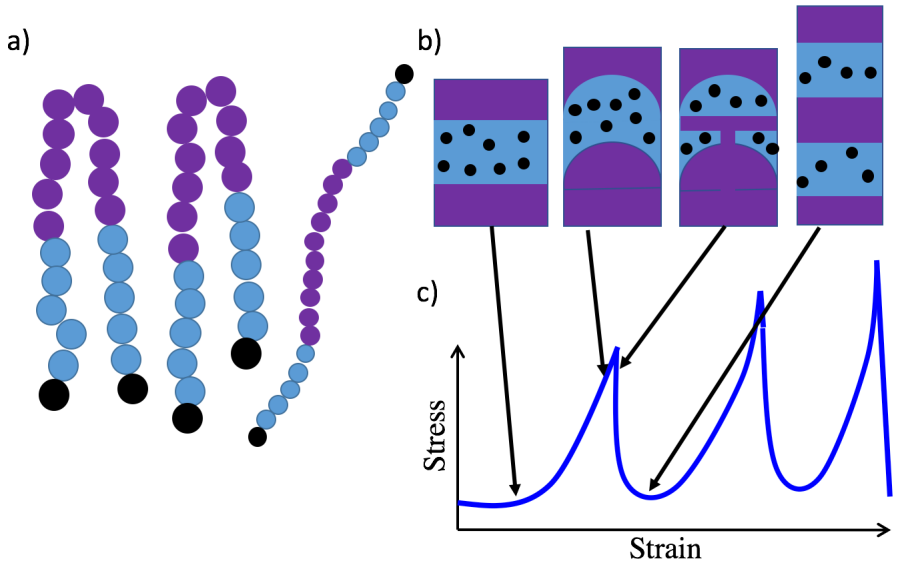

Figure 4.1: a) Depiction of unfolding mechanism for semi-flexible chains (colored as block copolymer for clarity). Unequal pull on the ends of the hairpin aids to the release of bending energy by unfolding. b) Cartoon of the stress built-up and release mechanism in T-LCE with block copolymer chains, describing states around a representative tooth in the stress-strain curve (c). *From left to right:* The applied strain first removes any "slack" in the system, and begins to deform the llamella, increasing the interfacial energy and the stress. After a greater amount of strain, the hairpins begin to unfold, breaking the parent layer into two daughter layers, and relaxing the stress.

In reaction schemes of flexible chain networks where the chains end-link with crosslinking molecules, different types of spatial/topological inhomogeneities develop; e.g., "self-biting" occurs when both ends of a chain react with the same cross-link. LCEs can exhibit additional microstructural defects like polydomains, which arise from a localized alignment of chains along distinct directors in different regions of the network, and degrade the mechanical properties [117,159,209]. In a preliminary simulation study of ABA block copolymer chains end-linked with tetrafunctional monomers, the saw-tooth tensile behavior found in T-LCEs was largely suppressed [5], likely due to the presence of defects and entanglements. With the goal of approaching ideal tetra-functional connectivity elastomers, a recently developed protocol uses two tetra-functional molecules that have either one of two reactive end-groups (amine terminated, or N-Hydroxysuccinimide terminated, depicted in Figure 4.2)[127,197]. Herein we refer to these tetra-functional molecules as monomers (type1 or type2, as depicted in Figure 4.1) and their arm lengths as l_a. Type1/2 monomers can only react with type2/1 monomers, preventing both self-biting and the formation of "loops" having an odd number of monomers, which improves mechanical properties [8]. A loop is defined as a connected path that starts from one given monomer and traces a path along connected monomers that ends at the starting monomer. n-loops are defined as loops that contain "n" monomers along the path. We will refer to this reaction scheme as the "AB tetra-monomer"(ABTAM) method.

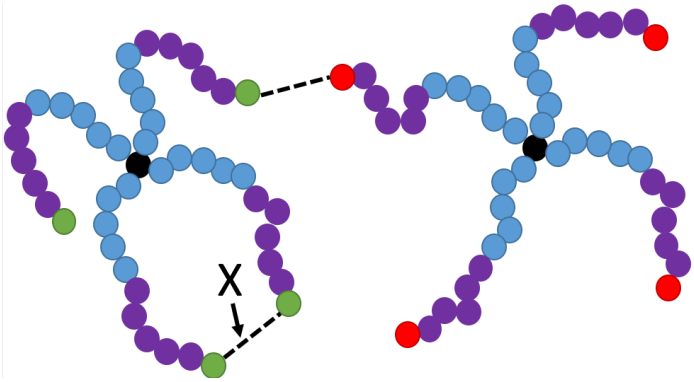

Figure 4.2: Schematic of ABTAM reaction. Each monomer consists of a tetra-functional cross-link (black), connected by the A-block (blue) of an AB block copolymer (B block is purple). There are two monomer types differentiated by the end-group on the chains (type 1 is red, type 2 is green). Only end-beads of different colors can bond.

Using molecular dynamics, this study aims to determine if realistic LCE networks, denoted as R-LCE, formed by the ABTAM protocol can exhibit the saw-tooth tensile response observed in T-LCEs, and if so, which synthesis parameters create networks with optimized toughness and α_m.

4.2 Models & Methods

Below we describe the key aspects of our simulation model/methods. A more complete description can be found in the supplementary information (SI). The polymers are described as bead-spring chains [71] with beads bonded using a FENE potential[70]:

$$U_{\mathrm{FENE}} = -K\frac{R_0^2}{2}\ln\left(1-\left(\frac{r}{R_0}\right)^2\right) + 4\varepsilon\left(\left(\frac{\sigma_{\mathrm{LJ}}}{r}\right)^{12} - \left(\frac{\sigma_{\mathrm{LJ}}}{r}\right)^{6} + \frac{1}{4}\right) \qquad (4.1)$$

where excluded volume interactions are described by the LJ-like potential term, that is cut off at $r=r_c=2^{1/6}\sigma_{\mathrm{LJ}}$. The spring constant is $K= 30\varepsilon/\sigma_{\mathrm{LJ}}^2$ and the maximum bond extension is $R_0 = 1.5\sigma_{\mathrm{LJ}}$, which describe a stiff bond whose equilibrium length is approximately equal to the bead diameter ($\sim 0.97\sigma_{\mathrm{LJ}}$). For non-bonded interactions, there are 4 unique bead types: A(B) beads represent the A(B) block, and C(D) represents the reactive end beads on monomers of

type 1(2). The tetra-functional bead is considered type A. A cut and shifted LJ potential is used with the form:

$$U_{\mathrm{LJ}} = 4\varepsilon^{ij}\left(\left(\frac{\sigma_{\mathrm{LJ}}}{r}\right)^{12} - \left(\frac{\sigma_{\mathrm{LJ}}}{r}\right)^{6} - \left(\frac{\sigma_{\mathrm{LJ}}}{r_c}\right)^{12} + \left(\frac{\sigma_{\mathrm{LJ}}}{r_c}\right)^{6}\right) \quad (4.2)$$

for $r \leq r_c$ ($U_{LJ} = 0$ otherwise), where $r_c^{ij} = 2.5\sigma_{\mathrm{LJ}}$ (attractive potential), when $i=j$, and $r_c^{ij} = 2^{1/6}\sigma_{\mathrm{LJ}}$ (repulsive potential), when $i \neq j$, and $\varepsilon^{ij} = 0.5\varepsilon$ for all interaction pairs ij. C and D beads interact through an additional Yukawa potential that is given in equation 4.8.

Chain stiffness is induced by a bending potential that is applied to all beads that are not bonded to a cross-link, which has the form:

$$U_{\mathrm{angle}} = k_b T K_{\mathrm{bend}}(1 + \cos\Psi) \quad (4.3)$$

where K_{bend} is a stiffness constant and ψ is the angle between two consecutive bond vectors; unless otherwise noted, $K_{\mathrm{bend}}=4$.

Structurally, the monomers are composed of a tetra-functional cross-link connected to 4 AB block copolymer "arms" by the A block. The last bead in each arm is treated as C/D beads for type 1/2 monomers. Arm lengths of l_a= 5, 11, 15, 20, 40, 60, 80 and 100 beads were studied, with corresponding volume fractions of A of 0.4, 0.455, 0.467, 0.5, 0.5, 0.5, 0.5, and 0.5. χN values for chains formed when two arms bond together are 13.5, 35.0, 52.7, 71.0, 161.5, 251, 342.8, and 434.6 in order of increasing l_a [112]. To simulate the cross-linking reaction, when a C/D bead is within $1\sigma_{\mathrm{LJ}}$ distance of a D/C bead and both beads have yet to form a cross-linking bond, a bond is formed with a preset probability. Following experimental studies that optimized the mechanical properties of synthesized networks [83,175], the two monomer types are at equal concentrations and the reaction rate is assumed to be kinetically limited (i.e. the probability of bond formation is low). Once initiated, the reaction proceeds

until conversion reaches ~96.5-97.5%. All results are for a good implicit solvent where both the A and B blocks of the monomers are well solvated, as would be the case if the solvent is a mix of 2 solvents, each selective for one block. Standard periodic boundary conditions are used throughout. Once synthesized, the networks are compressed from the synthesis concentration, ϕ_0, to the prescribed melt density, ϕ_{melt}, and equilibrated [ϕ is defined as (number of beads)/σ_{LJ}^3]. For uniaxial deformation, we use an "iso-strain isobaric mixed ensemble" where the deformation axis has a prescribed strain while the orthogonal directions are controlled by a barostat [144].

4.3 Results and Discussion

4.3.1 Effect of Monomer Concentration at Crosslinking

Our simulations aim to quantify how network quality is affected by ϕ_0, the monomer concentration during crosslinking. ϕ_0 has a strong effect on the topological microstructure of the network: a small value can lead to inhomogeneous, poorly percolating networks, while a large value can lead to highly entangled systems. Synthesized networks of varying l_a and monomer concentration all microphase segregate into AB lamellae when compressed to $\phi_{melt,}$ regardless of ϕ_0; however, the lamellae plane normal generally doesn't point in the [100] direction due to a mismatch between the lamellar spacing and the simulation box size [219]. This leads to the deformation axis being initially misaligned with the director (the direction of chain alignment), which leads to "soft" low-stress initial deformations (see SI)[129]. Figure 4.3 shows the tensile responses of several networks synthesized at different concentrations for l_a = 11 (Figure 4.3a) and 20 (Figure 4.3b).

After the soft regime, all curves approximately follow a characteristic pattern with ϕ_0 determining α_m. The toughness, $\gamma_{\alpha=x}$, is defined as the area under the stress strain curve from α=1 to x (omitting negative stresses). We find that $\gamma_{\alpha=5}\sim$ 2.43 for the R-LCE network with l_a =11 and ϕ_0=0.045, comparable to $\gamma_{\alpha=5}$ ~1.91 for T-LCE. The microscopic mechanism underlying the saw-tooth behavior is also fundamentally the same as that of the perfect network (summarized in the introduction). The behavior changes when the stress starts to diverge, where, due to entanglements, enthalpic segregation of blocks and chain relaxation are hindered. At that point, chain disinterdispersion, hairpin unfolding and phase segregation subside, and the tensile response becomes dominated by chain hyperextension and bond stretching. Indeed, near α_m, bond-stretch energy trends upward while interfacial energy is not a significant contributor to the stress as shown in Figure 4.10 in the SI. At very low ϕ_0 (ϕ<0.01 for l_a=11) the connectivity of the network is significantly reduced compared to higher ϕ_0 cases (ϕ_0>0.1 for l_a=11) and may be affected by the finite size imposed by the boundary conditions. To clarify this effect, we simulated a network synthesized with 4096 monomers (8 times the 512 used for all other syntheses) at ϕ_0=0.0133 and saw similar connectivity (Figure 4.10). The tensile response of this network (Figure 4.3a curve labeled "large" and Figure 4.3c) shows peaks with smaller and less defined peaks. This is due to the longer relaxation time scales in larger networks, where the averaging of multiple transformations initiated at slightly different times leads to a smoothing out of the tensile response. This smoothing is expected to continue into the macroscopic regime, especially above a length scale where regions of different chain alignment form (known as polydomains)[209]. These domains will only be weakly coupled to each other at their

boundaries, leading to further de-coherence of the formation of new smectic domains in the sample, and to additional "averaging" of the tensile response.

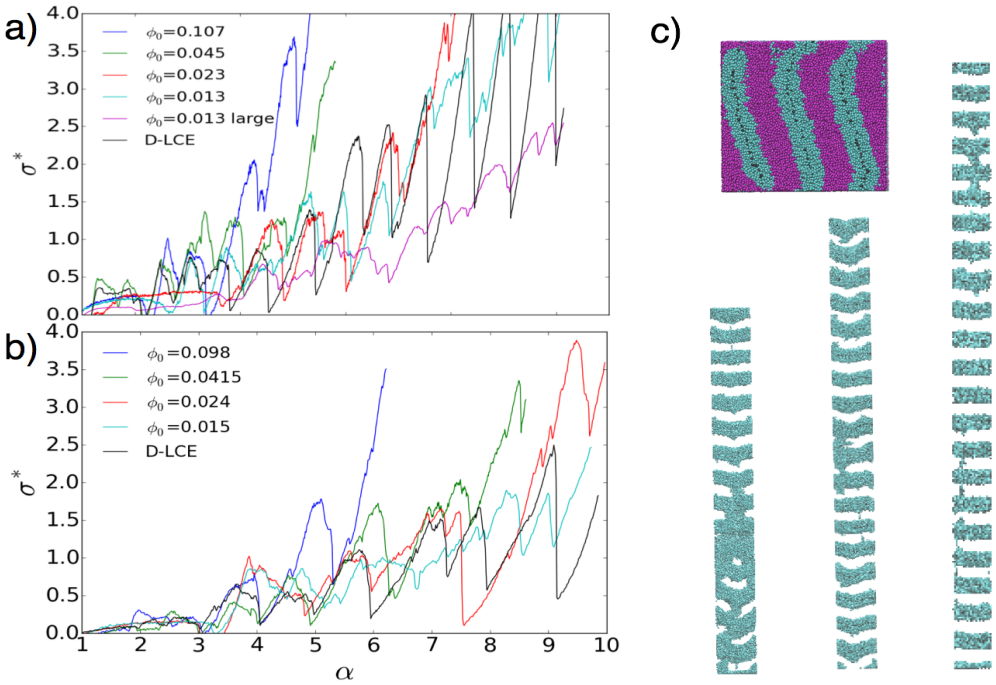

Figure 4.3: Tensile plots of various ϕ_0 for l_a=11(a), and l_a=20(b). (c) Snapshots from the deformation of a network synthesized from 4096 monomers at ϕ_0=0.0133, for α=1, 4.7, 5.9, and 6.9 counter-clockwise from the top left. Color scheme follows Figure 4.1. For high α the B block is removed for clarity.

Having established that the tensile response of R-LCEs approaches that of the T-LCE, we now examine how the microstructural details affect the tensile response. Figure 4.4a shows the number of 2,4,6-loops normalized by the number of monomers. Figure 4.4b shows the equilibrium swelling volumes normalized to the swelling volume of a perfect network (swelling ratio = *SR*) and the chain ratio (CR) [153] (see SI for details). *CR* compares the ensemble average of the minimum length a chain can have while still obeying topological entanglements (found by shrinking, from their fixed crosslinked end points, all chains till taut), to the minimum distance between the chain end points; thus, larger values of *CR*

represent more entangled networks. SR provides a more macroscopic indicator of network quality, with smaller values pointing to more trapped entanglements.

Figure 4.4a and Figure 4.4b show data for all arm lengths plotted against half the average monomer separation distance at ϕ_0, r_{sep} (based on nearest neighbors from Voronoi tessellation), normalized by the average end-to-end distance of individual chain arms, r_{ee}. The ratio r_{sep}/r_{ee} represents the separation distance relative to the distance where the monomers begin to "feel" each other and $r_{sep}/ r_{ee}\approx 1$ can thus be seen as a probe of conditions consistent with the "overlap" concentration, ϕ^*. At $r_{sep}/r_{ee}\approx 0.9$, there is an inflection at $CR \sim 1.05$, corresponding to $SR \approx 1$, at which point the number of 6-loops start decreasing with decreasing ϕ_0, signaling a change in microstructure.

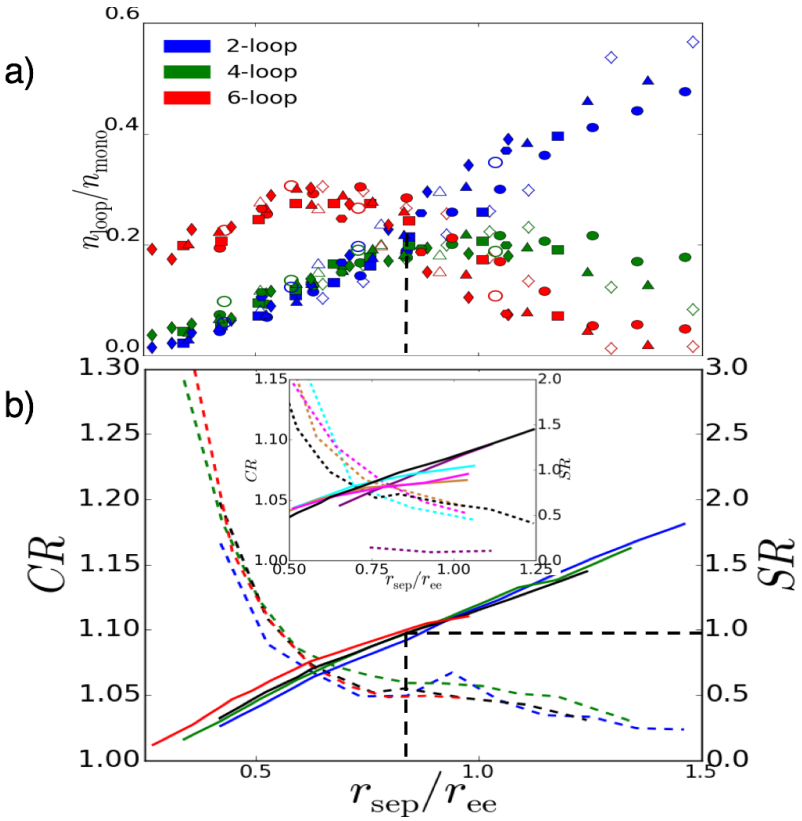

Figure 4.4: The number of topological loops (a), *SR* (b, solid lines), and *CR* (b, dashed lines), for different l_a. In (a), l_a =5, 11, 15, 20, 40, 60, 80, and 100 are represented by the open diamonds, filled circles, filled squares, filled triangles, filled diamonds, filled hexagons, open circles, and open triangles respectively. In (b), l_a =11, 15, 20, and 40 are colored blue, green, black, and red, respectively. In the inset, l_a= 5, 20 (for reference), 60, 80, and 100 are colored purple, black, cyan, brown, and pink, respectively. Black dashed lines outside of inset are used as guides to show that when *SR* = 1, there is an inflection point in *CR*.

For $11 \leq l_a \leq 40$, two regimes of network topology as a function of ϕ_0 are postulated to explain these observations. The first topological regime corresponds to $r_{sep}/r_{ee} < 0.9$ (Figure 4.4a and Figure 4.5a), where increasing r_{sep}/r_{ee} (or decreasing ϕ_0) decreases the number of trapped entanglements, n_e, because of the reduction in chain interpenetration during synthesis. Reducing n_e directly lowers both the *CR* and the effective cross-link density which translate into larger swelling. Also in this regime, the number of loops of any size decreases with decreasing r_{sep}/r_{ee} because higher concentrations increase the number of neighboring

monomers, which reduces the probability of creating loops of any size. These trends continue until $r_{sep}/r_{ee}\approx$ 0.9 (Figure 4.5b), where the number of 6-loops is maximized, the $SR\approx1$, and the CR curve has an inflection.

The second topological regime corresponds to $r_{sep}/r_{ee}>0.9$ (Figure 4.5c), where a fully connected network would result in chains that are over-stretched and so the network collapses to allow chains to relax. In this regime, α_m is directly correlated to the number of 2-loops and inversely related to the number of 6-loops. If the probability of a loop becoming entangled is proportional to the loop "volume", 4,6-loops are more likely to become entangled compared to 2-loops, and are consequently more deleterious to a saw-tooth, large α_m tensile response. Since larger loops are areas of stress localization during deformation[127], both the increase in 2-loops and decrease in 4,6-loops lead to larger α_m. The inset of Figure 4.4b shows that for l_a=5 and l_a=60-100 deviations from the two-regime trends discussed above become significant. Indeed, very short chains (of the order of the persistence length) entangle weakly over the range of ϕ_0 considered, while chains longer than some threshold [198] entangle more strongly for $r_{sep}/r_{ee}<0.9$ than shorter chains. Henceforth, we only focus on systems with $5<l_a<60$.

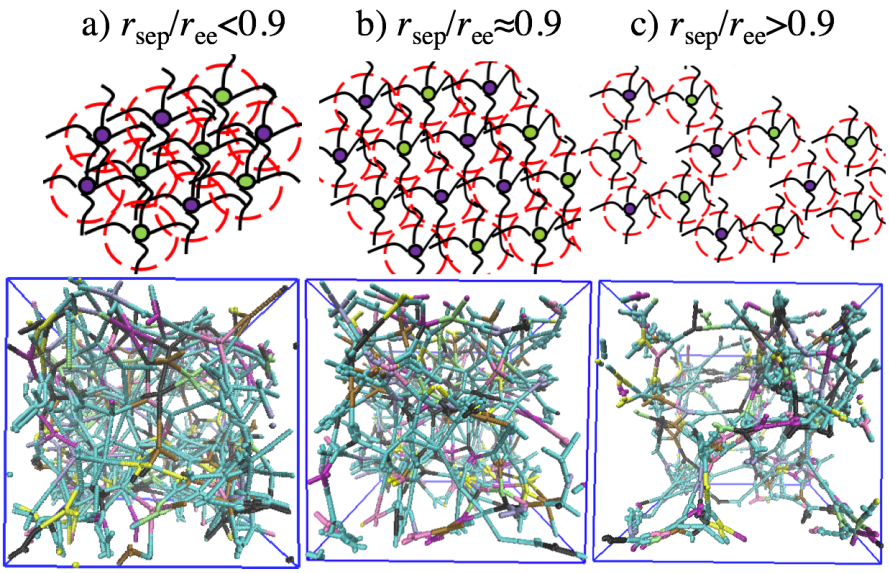

Figure 4.5: Cartoons (top) of synthesis environment (red circles depict monomer coils of radius r_{ee}), and representative simulation snapshots after applying the chain ratio algorithm (bottom) for networks with l_a= 11 synthesized at ϕ_0= 0.107(a), 0.045(b), and 0.023(c).

To define an optimal ϕ_0, in Figure 4.6 we plot α_m scaled by $\alpha_m^{ideal} = \alpha_m$ of a T-LCE with the same l_a, against r_{sep} scaled by r_{ee}. The plot is linear but plateaus for $r_{sep}/r_{ee} \geq 1$. These trends can be rationalized using a simple physical picture, by imagining the monomer coils as spheres of radius r_{ee}, and the overlap volume between spheres, V_{cap}, as dictating n_e and consequently α_m. Since V_{cap} is largely linearly dependent on r_{sep} until plateauing at zero when $r_{sep} \geq r_g$, then the assumption that $n_e \propto V_{cap}$ provides predictions consistent with both the linear regime (r_{sep}/r_{ee}<1) and the plateau regime (because for $r_{sep}/r_{ee} \geq 1$ there is zero overlap and n_e no longer decreases). As a result, R-LCEs synthesized at $r_{sep}/r_{ee} \sim 0.9$ are predicted to have the greatest toughness, by maximizing α_m, while still having an isotropic network connectivity. Such optimal conditions approximately correspond to ϕ_0= 0.023 and 0.012 for the networks with l_a= 11 and 20 respectively (see Figure 4.2 for stress response and Figure 4.10 in the SI for representative snapshots).

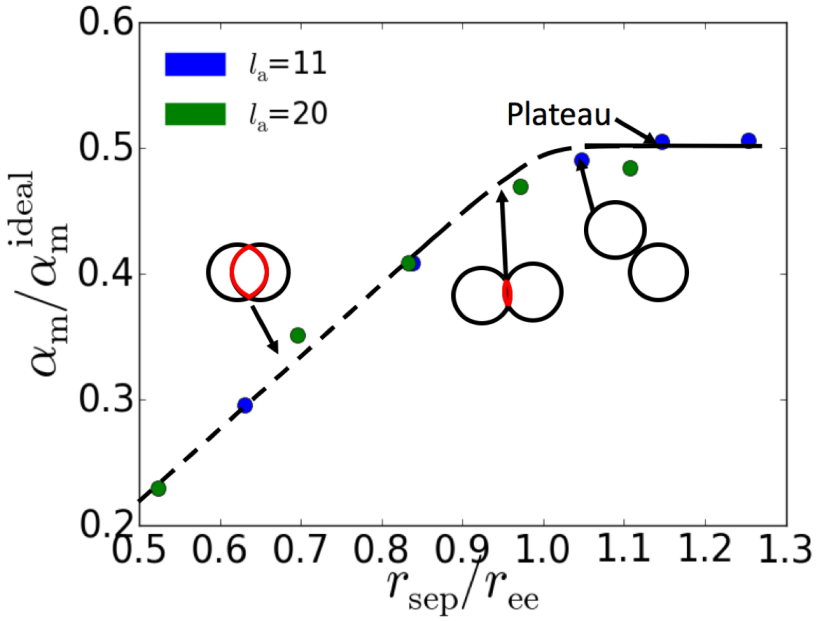

Figure 4.6: Plot of the scaled extensibility against the r_{sep} normalized by r_{ee}. Circle cartoons represent the monomer spheres and their interacting regions (spherical caps in red).

The trends in Figure 4.6 can also be explained by analyzing the behavior of the free energy "*F*" for an entangled network, according to the theory developed by Edwards, and Vilgis [165]:

$$\frac{F}{k_BT} = \frac{N_c}{2}\left\{\frac{(1-a^2)\sum_{i=1}^{3}\alpha_i^2}{1-a^2\sum_{i=1}^{3}\alpha_i^2}\right\} + \frac{N_s}{2}\left\{\sum_{i=1}^{3}\left[\frac{\alpha_i^2(1+\eta)(1-a^2)}{(1+\eta\alpha_i^2)(1-a^2\sum_{i=1}^{3}\alpha_i^2)} + \ln(1+\eta\alpha_i^2)\right] + \left(1 - {}^{N_c}/_{N_s}\right)\ln(1-a^2\sum_{i=1}^{3}\alpha_i^2)\right\} \quad (4.4)$$

where a is the ratio of the length scale of a polymer random walk to the length scale of the primitive path, N_c is the number of cross-links, N_s is the number of slip links, η is a measure of slippage, and α_i is the extension ratio of direction i. The criterion for maximum extensibility is when $dF/d(\alpha_x^2)$ diverges (taking x as the extensional direction). By taking the derivate of F with respect to α_x^2 and applying the constraint that α_x is large, we get:

$$\frac{dF}{d(\alpha_x^2)} \propto \frac{N_c}{2}\left\{\frac{(1-a^2)}{1-a^2\alpha^2} - \frac{(1-a^2)a^2\alpha^2}{(1-a^2\alpha^2)^2} - \frac{a^2}{1-a^2\alpha^2}\right\} + \frac{N_s}{2}\left\{(1+\eta)(1-a^2)\left[\frac{a^2}{\eta(1-a^2\alpha^2)}\right] - \frac{a^2}{1-a^2\alpha^2}\right\} \quad (4.5)$$

Eq. 4.5 has two limits of interest: $a\sim1$ (high entanglement), and $a\sim0$ (no entanglement).

The limit, $a \sim 0$ (no entanglement, large r_{sep}/r_{ee}), leads to:

$$\frac{dF}{d(\alpha_x^2)} \sim \text{constant} \qquad (4.6)$$

which predicts the plateau observed in our results. For the $a\sim1$, the expression simplifies to:

$$\frac{dF}{d(\alpha_x^2)} \propto \frac{1}{\alpha^2 a^2} \quad (4.7)$$

To maintain the same value of the derivative (i.e., the divergence slope for maximum extensibility) the product $a^2\alpha^2$ must remain the same, so $\alpha\propto1/a$. Parameter a is inversely proportional to *CR*, and since Figure 4.4b also showed that *CR* is roughly inversely proportional to r_{sep}/r_{ee} for $r_{sep}/r_{ee}<0.9$, it follows that $\alpha_m\propto r_{sep}/r_{ee}$ (consistent with Figure 4.6).

4.3.2 Effect of Chain Flexibility

Chain backbone stiffness is another key design parameter that affects the topological and tensile properties of the networks prepared using the ABTAM scheme. Since experimental studies have mostly focused on flexible polymers, we also simulated fully flexible chains. Our results for such systems indicate that the data for number of loops, *SR*, and *CR* all collapse onto similar curves as those reported before for semi-flexible chains when using the r_{sep}/r_{ee} scaling. Unexpectedly, however, we also find that fully-flexible chains cannot reproduce the distinctive saw-tooth behavior seen for semiflexible chains (Figure 4.7).

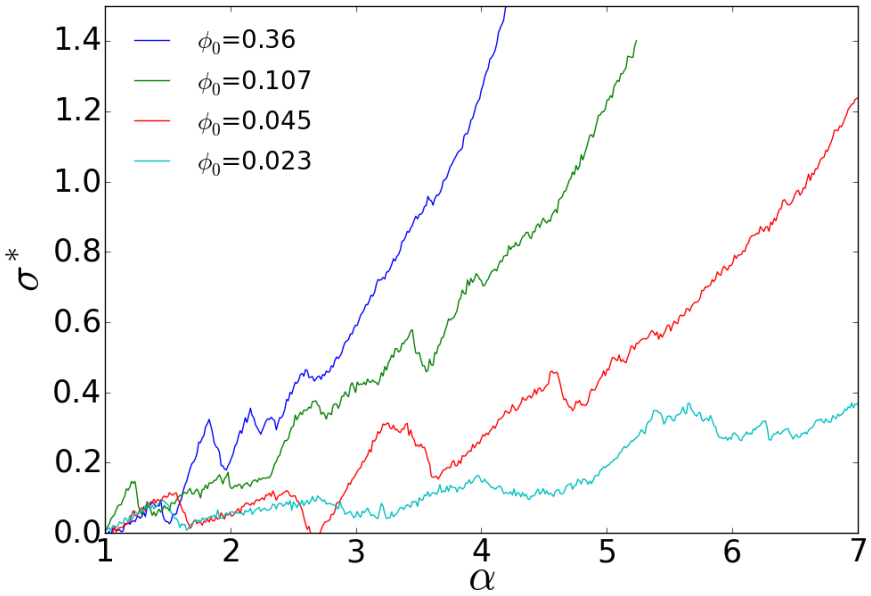

Figure 4.7: The tensile response of synthesized flexible block copolymer networks for different ϕ_0.

To decouple the effects of chain backbone and network topology on the tensile response, we prepared one network with "flexible topology" (taken from a synthesis where the chains were fully flexible) that was equilibrated and deformed in two different scenarios: one where the chains were still fully flexible (to be denoted the FF network), and one where the chains were now equilibrated as semi-flexible (to be denoted the ESF network). Figure 4.8a shows the tensile response of the ESF network, exhibiting a saw-tooth behavior which is qualitatively similar to the tensile response of a network with semi-flexible chains from synthesis to deformation, shown by the line marked "SF" (the values of r_{sep}/r_g were matched for the two networks). This comparison indicates that the topology of fully flexible networks is not responsible for the lack of the saw-tooth behavior.

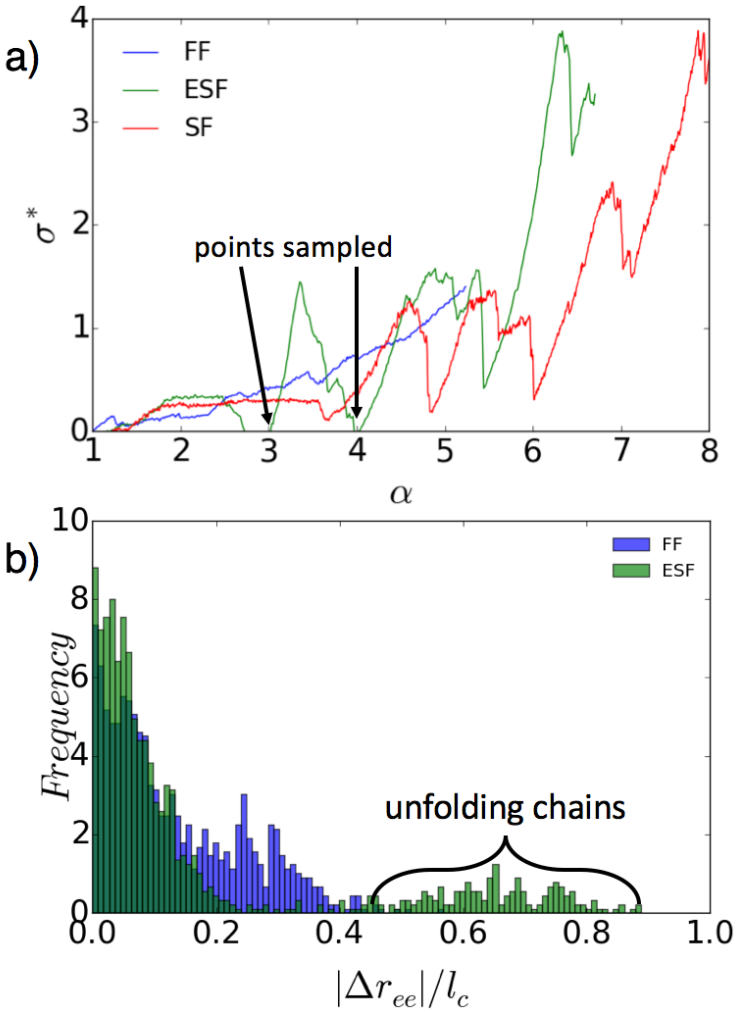

Figure 4.8: a) The tensile response of FF, ESF, and SF networks. "FF" and "ESF" ("Equilibrated as SF") have identical network topologies. Note: the tensile response of the FF network is the same as that shown in Figure 4.6 (for ϕ_0=0.107). b) Plot of the absolute change in r_{ee} normalized by l_c for all chains in the "FF" and "ESF" networks. The two values of α from which the configurations were sampled are marked by arrow in panel a).

To identify the key microstructural difference between the FF and ESF networks, we calculate the instantaneous root mean squared chain end-to-end distance of the full-length chains in the network. Using representative configurations from the deformations of the FF, and ESF networks, the value of r_{ee} was calculated for each chain at two values of α, corresponding to before and after a "tooth" in the tensile response in ESF (as marked in Figure 4.8a). In Figure 4.8b we plot the distribution of the absolute change in r_{ee} normalized

by chain contour length, $|\Delta r_{ee}|/l_c$, for both networks. The distribution for the FF network is largely unimodal while that for the ESF distribution has, besides a main unimodal chain population exhibiting small-to-moderate deformation, a subpopulation of chains undergoing very large unfolding, indicative of a less affine deformation. The degree of affine deformation is attributed to the way in which chains pack when they phase segregate. For the ESF/SF networks, the chains orientationally align (as expected of mesogenic chains) which leads to the formation of hairpins in the network. Mobility in the melt along the deformation direction is required to unfold the chains, which is augmented by the nematic/smectic phases seen in T/R-LCE, where diffusion/mobility is greater along the director [52]. Orientational order also favors the complete unfolding of hairpins; in fact, the deformation predominantly occurs only where the new domain forms (associated with the population of chains with $|\Delta r_{ee}|/l_c > \sim 0.4$ in Figure 4.8b). In contrast, flexible chains do not orientationally order and any unfolding of hairpins will likely encounter obstructing chains and preclude localized deformations. Therefore, a fully-flexible network undergoes more affine deformations (Figure 4.7b) as compared to semi-flexible networks (Figure 4.8a). Additionally, because there are no "true" hairpins in the FF network, the mechanism of unfolding and new domain formation does not give "extra extension" to the network nor releases the stress abruptly, both key ingredientes for a saw-tooth tensile response.

4.4 Summary & Conclusions

Unlike previous studies that detected saw-tooth responses with idealistic network topologies, this work provides clear guidelines for engineering realistic, super-tough elastomers, practicable via ABTAM-like approaches. Our R-LCEs achieve maximum

extensibilities of α_m=5, and 8, and toughness values in the order of $\gamma_{\alpha=5}$= 0.96 GPa, and 0.52 GPa (assuming one unit of toughness is equal to 0.4GPa [217]), for l_a=11, 20, respectively. Such values compare favorably to spider-dragline silk (α_{break}=1.6 [148] and toughness ~ 0.15 GPa) and networks of PEO chains using ABTAM (α_{break}=1.9, and toughness~ 0.00023GPa[175]).

The quality of the network topology produced by the ABTAM scheme depends crucially on the concentration during crosslinking ϕ_0. We have identified a microscopic metric that provides a tighter correlation with the network properties than ϕ_0: the average separation radius between monomers scaled by the arm end-to-end distance, r_{sep}/r_{ee}. Our results suggest that networks synthesized at $r_{sep}/r_{ee} \approx 0.9$ provide the best combination of extensibility, α_m, and "quality" of the saw-tooth tensile response. When r_{sep}/r_{ee} is increased, entanglements decrease and α_m increases, however, for r_{sep}/r_{ee}> 0.9, the reduction in entanglements abates and α_m plateaus.

We find that a two-regime trend for topological "quality" (Figure 4.4) hold true for $5<l_a<60$ (which encompasses experimental ABTAM syntheses [175]; e.g., degree of polymerization ~50, equivalent to l_a~25), for chains with persistence length of ~4 (capable of accommodating liquid crystallinity) and for fully flexible chains. The local ordering and deformation mechanism of semi-flexible chains leads to the non-affine network deformations that are necessary for the saw-tooth tensile response to occur.

4.5 Acknowledgements

This work was supported by the National Science Foundation Award CMMI 1435852. This

work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1053575.

4.6 Supplementary Information

4.6.1 Simulation model and methods

Molecular dynamic runs were conducted using LAMMPS [163] and systems visualized using VMD [87]. During the crosslinking process, the reactive end groups, C and D, carry charges. The crosslinking reaction rate is experimentally controlled by changing the ionic strength of the buffers used, which alters the long-range interactions among the reactive end groups. Hence, on top of the LJ potential between end-groups (a repulsive potential), the electrostatic interactions are accounted for by imposing a Yukawa potential:

$$U_{\text{yukawa}} = A^{ij}\, e^{-\kappa r}/r \qquad (4.8)$$

for $r/\sigma_{LJ} \leq 5$ ($U_{\text{yukawa}} = 0$ otherwise), where A^{ij} is a proportionality constant, and κ is the screening length. We set $A^{ij}=25\varepsilon$ if $i=j$, $A^{ij}=-25\varepsilon$ if $i\neq j$, and $\kappa = 2$ for $i,j = 3,4$, choices aimed at lowering the reaction rate. For values of $|A^{ij}|>50$, there is evidence of the end-groups clustering into primitive cubic arrays. Thus, we choose $|A^{ij}|=25$ to mimic a buffer with high ionic strength. When beads are directly bonded to each other, the calculation of non-bonded potential between them is omitted. Reported properties are scaled as follows: $T^*=Tk_b/\varepsilon$, $\tau=t(\varepsilon/(m\sigma_{LJ}^2))^{1/2}$, $\sigma^*=\sigma\sigma_{LJ}^3/\varepsilon$, $P^*=P\sigma_{LJ}^3/\varepsilon$, $\gamma^*=\Gamma\sigma_{LJ}^3/\varepsilon$, and ρ = (# of particles)/ σ_{LJ}^3, where T is temperature, σ is stress, P is pressure, Γ is the toughness, ϱ is number density, and all scaling variables (σ_{LJ}, m, ε) are set to 1.

Initially, two types of reaction simulations were run, which consider the effect of the quality of the implicit solvent on the system. For “single-solvent” syntheses, where the

solvent is only selective for the B block, T^*=2.0. For "ideal" synthesis simulations, where both blocks are well solvated to mimic two solvents, each selective for one of the blocks, (i.e. the cutoff radii are set to the LJ minimum r_c^{AA}= r_c^{BB}= $2^{1/6}$), and T^*= 1.0. In all synthesis simulations, Langevin dynamics are applied with a damping factor =1τ. For melt compression, equilibration, and tensile tests, T^*= 2.0 with a Nosé-Hoover thermostat. The equation of motions for all simulations are integrated using a velocity Verlet algorithm, with a timestep of .005τ.

To form new bonds during the simulation, the distance between all pairs of C and D beads are evaluated every 10 timesteps, and if any of these distances is ≤ σ_{LJ}, a bond is created with probability between 0.01 and 1. Once reacted, the C and D beads become type B beads. When comparing the "ideal" two-solvent synthesis with a one-solvent synthesis, metrics of network quality like the number of loops, swelling ratio, and chain ratio, are similar between both reaction schemes, and indicating that the effect of solvent quality is negligible because the monomer concentrations are too dilute to cause any substantial aggregation of the monomers. Results presented in the main text are for networks from "ideal" syntheses.

As described in the main text, for very low densities, the equilibrium swelling density of the gel is greater than the synthesis density. Thus, in a realistic synthesis, the gel would contract, effectively increasing the synthesis density. To more closely simulate the behavior of these networks, during an "ideal" synthesis simulation at T^*=2.0, the pressure is measured as a function of the conversion, and a critical conversion value of ~67% is observed for P^*=0 (the condition for equilibrium swelling with implicit solvent). When the network reaches this conversion, the simulation is set to be isobaric (NPT ensemble), where a Berendsen barostat is used to control the simulation box dimensions. At reaction completion, the

density of the network can increase upwards of 100%. The results indicate no large difference between this NVT-NPT hybrid approach and the purely NVT ensemble approach. The only notable difference between the two methods is that the hybrid approach reaches the final conversion more quickly, due to the gradual density increase. Results in the main text are for NVT reactions only for their higher computational efficiency of NVT and for circumventing the need to calculate the critical conversion where P^*=0.

To achieve equilibrated melts, ε^{ij} is set to 0.25ε, and the system is compressed in two stages: a compression to a density of ρ=0.5 in 10^7 time steps, followed by a slow compression to ρ =0.85 in 5×10^7 time steps. ρ =0.85 is chosen to be the melt density as the bead-spring model matches real polymer melt behavior at this ρ[70]. The compression follows an exponential form, $l=l_0\exp(k\tau)$, where l is the simulation box length, l_0 is the original box length, k is a dimensionless compression rate, and τ is the scaled LJ time. The slow compression from ρ=0.5 to 0.85 allows for enhanced equilibration as chains can rearrange more easily.

Once the melt density is reached, ε^{ij} is set to 0.5 and the network equilibrates until steady state values are attained for the bending, non-bonded, and bonded energies, as well as the diagonal pressure tensor components, p_{xx}, p_{yy}, and p_{zz}. The time scale of equilibration is $\sim 5\times10^7$ time steps. This state is considered the undeformed state, α=1, where α is the extension ratio, i.e., the length of the longest box edge scaled by the undeformed cubic box length (l_{ext}/l_0). For uniaxial deformations, the original lamellae orientation dictated how the box was deformed. For plane normals not pointing in the [1 0 0] direction, three tests were performed, deforming the network in all three orthogonal directions. For plane normals pointing in the [1 0 0] direction, three replicate tests for that configuration (reassigning new

random velocities and equilibrating before the start of each replicate) were performed along the [1 0 0] direction. During uniaxial extension, the box length along the deformation axis (x, for example) is expanded linearly every 5000 time steps up to α=11 in 2×10^8 time steps, leading to a deformation rate of $10^{-5}\ \alpha/\tau$. The orthogonal directions (y, and z) are controlled by a Nosé-Hoover barostat, set to couple the two orthogonal directions so the cross-section maintains the same aspect ratio (this method is referred to as the Mixed Ensemble method, ME, in the main text). This procedure is preferred because in purely constant-volume simulations, chains can pack into a crystalline arrangement and get "stuck" because the cross-sectional area is never allowed to expand to allow chains to leave the crystal. The pressure values used for the barostat are taken from the orthogonal pressure tensor components (P_{yy} and P_{zz}) of networks that equilibrate with lamellar plane normal pointing in the [1 0 0] direction. During ME deformation, the density of these networks fluctuates with a systematic increase at higher strains as seen in experimental systems[217], however, the density roughly stays around the prescribed density ρ=0.85. The stress during deformation is calculated using:

$$\sigma^* = -P^*_{ii} + \lambda\left(P^*_{jj} + P^*_{kk}\right) \qquad (4.9)$$

where P_{ii}, P_{jj} and P_{kk}, and λ are the extensional direction pressure tensor component, the orthogonal pressure tensor components, and the Poisson ratio, respectively. The average λ is calculated for all ME tensile tests while $\lambda\cong0.5$ for all constant-volume tests. Tensile plots presented in the main text give the stress assuming this averaged value of λ rather than using the instantaneous value of λ at each α.

The toughness, γ, or energy absorbed during deformation, is calculated using:

$$\gamma(\alpha_f) = \int_1^{\alpha_f} \sigma^*(\alpha)\, H(\sigma^*) \mathrm{d}\alpha \quad (4.10)$$

We are interested in the total energy absorbed (i.e., total work done on the system), so $H(\sigma^*)$, the Heaviside function, is introduced to ignore negative stresses. We thus neglect the energy "returned" (work done by the system) when stress is negative which would be dissipated. Since our model bonds do not account for breaking events (future work will explore breakable bonds), there is no value of α_f where the network breaks, so we compare values of $\gamma(\alpha_f)$ where α_f is below the point where the stress diverges (for most systems, $\alpha_f \sim 5$). To provide some calibration for the toughness values to be reported, we can compare the simulated toughness of ~ 0.15GPa[46] (calculated for $1 < \alpha < 5$) for an atomistic model of a polyethylene network with diamond connectivity to γ=0.33 for a fully flexible coarse-grained homopolymer network of comparable molecular weight; hence a unit in γ in our model would correspond to ~0.4GPa.

Repeated runs of the same network (same synthesis concentration) give slightly differing results; accordingly, the plots presented are intended to be representative and whenever possible average values are reported. Quantities varying with α are plotted showing values block averaged over intervals of ~0.005-0.01α.

4.6.2 Calculations for Network Characterization

4.6.2.1 Swelling Ratio

To calculate the swelling ratio (*SR*) and only consider the effects of network topology only, the network is simulated with $r_c^{i,j}=2^{1/6}$ for all pairs *i,j*, and a barostat is used to set the pressure to zero, allowing each dimension to fluctuate uncoupled from the other directions. The ensemble average of the swollen simulation box density is recorded and is normalized by the swollen volume of a perfect, unentangled diamond network (therefore by definition, *SR* for a perfect network is 1).

4.6.2.2 Chain Ratio

The chain ratio (*CR*) calculation is based on previous methods used to quantify the primitive path length of an entangled chain [199]. In summary, each monomer "arm" is assigned a unique type, giving a number of types equal to 4 times the number of tetrafunctional "crosslinks". Cross-links are all assigned to another type and kept at fixed positions throughout. All types are mutually repulsive to each other (i.e. $r_c^{i,j}=2^{1/6}$, $\varepsilon^{i,j}=1.0$ for all i,j) except for the type pairs who are bonded to each other where $\varepsilon^{i,j}=0$ (i.e., if the arms with type x beads is bonded to the arm with beads of type y, $\varepsilon^{x,y}=0$). Additionally, the parameters K and ε in Eq. 4.1 are set to 100 and 0, respectively; these parameter choices allow the chains to contract, shortening their length due to the larger spring constant and lack of repulsion between beads along the same chain. Since entanglements are the only thing impeding each chain from achieving a contour length equal to the minimum distance between the two terminating crosslinks, the ratio of the equilibrium contour length to the minimum contour length, l_{eq}/l_{min}, is the definition of the chain ratio, and provides a measure of how entangled the network is.

4.6.2.3 Number of loops in synthesized network topologies

An important metric for charactering network quality is the number of n-loops, where n=2,4,6. The list of bonds created during the simulation are converted into connections between monomers, which are treated as a group of nodes (the monomers), connected by edges (bonds formed during the synthesis). A loop search algorithm is applied where paths are drawn from a starting node to any node up to 3 edge connections away (see Figure 4.9).

If two paths do not share any nodes (except for the first and last nodes on the path), and the paths end at the same node, these paths form a n-loop, where n is the sum of the number of edges along both paths. If a node is connected to another node more than once, it is classified as a 2-loop. If a node is connected by more than 2 bonds to another node it is still classified as a single 2-loop; however, the prevalence of such structures is very low and often negligible. For a single node, only loops of the smallest n are considered, i.e., if a 4-loop is the smallest loop found, 6-loops are neglected. This process is repeated for all monomers, and only unique loops are counted. Then number of n-loops is determined from counting loops of length n in the loop list.

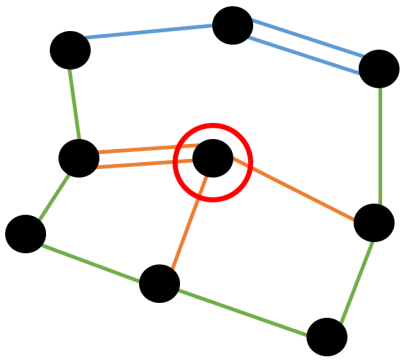

Figure 4.9: Example of loop calculation for monomer circled in red. Paths which are 1, 2, and 3 edge lengths away are colored orange, green, and blue respectively. This monomer is part of a 2-loop, 2 4-loops, and a 6-loop, but the 4- and 6-loops are ignored because the 2-loop is the smallest.

4.6.3 Additional Supporting Results

4.6.3.1 Comparison between perfect diamond (D-LCE) and realistic network (R-LCE)

Figure 4.10compares snapshots from the deformation of R-LCE at optimal conditions (l_a=11, ϕ_0=0.0133, r_{sep}/r_{ee}=1.25) to those corresponding to D-LCE of equal l_a. The morphology of the networks match until very large α. At large α, the entanglements restrict the R-LCE cross-

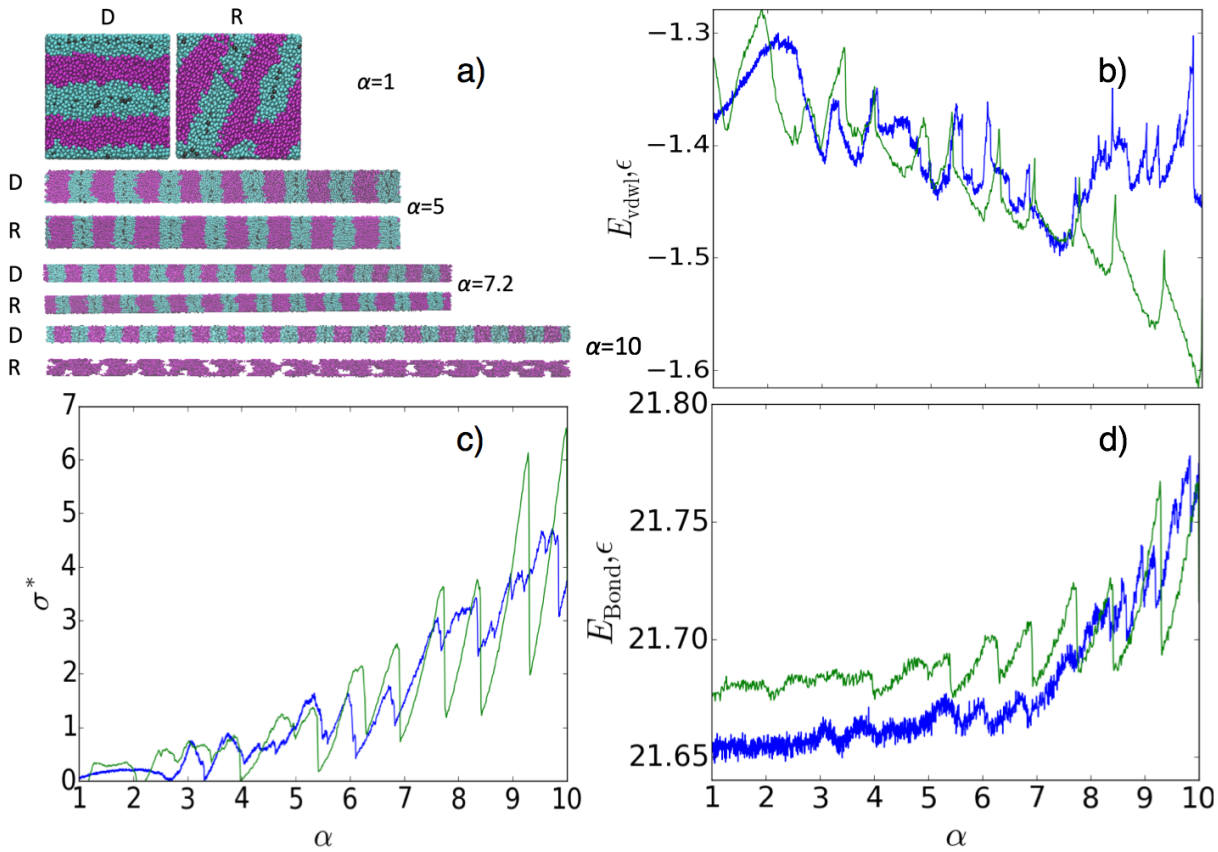

Figure 4.10: a) Snapshots of D-LCE (D) and R-LCE (R) for various values of α. A block, B block, and cross-links are colored as blue, purple, and black, respectively. For α=10, the B block is removed from the snapshot for R-LCE to show the effects of the entanglements. Example non-bonded energy, E_{vdwl}(a), tensile response (b), and bonded energy, E_{Bond} (c), for a R-LCE (l_a=11, ϕ_0=0.0133, $r_{\mathrm{sep}}/r_{\mathrm{ee}}$=1.25, blue lines), and a D-LCE (green lines).

link mobility. Upon stress divergence (i.e., when the stress stops dropping down significantly), at α~6.5 (Figure 4.10c), the non-bonded energy, E_{vdwl} (Figure 4.10b), decreases, and E_{Bond}, the bonded energy (Figure 4.10d), increases. E_{vdwl} has two contributions, namely, from 1) interfacial energy (from the unlike-bead interactions), and 2) excluded volume interactions (from the repulsion of beads being "squeezed" together). The relative contribution of the latter can be estimated from E_{Bond} by comparing D-LCE (green lines) and R-LCE (blue lines). We note that E_{Bond} for D-LCE is higher than for R-LCE at α=1, as reflected by the higher value of E_{vdwl} for D-LCE even though its initial morphology has a lower surface area, which would result in a smaller contribution to E_{vdwl} from interfacial energy.

For α>6.5, E_{vdwl} for R-LCE increases but tends to plateaus, while E_{Bond} continues to increase. These observations indicate that the primary cause for the rise in stress after a critical deformation is the effect of entanglements restricting chain motion.

4.6.3.2 Finite size effects

To verify that finite size effects were not playing a large role in the connectivity of the network while the crosslinking process is taking place, a synthesis was carried out using 4096 monomers (8 times larger than the syntheses shown in the main text) with arm length of 11 beads. The resulting network structure, shown in Figure 4.11, displays no signs of weaker connectivity, and all the metrics for network quality show numbers commensurate with those of the smaller system. Additionally, Figure 4.11 shows the tensile response of this network, which is also similar. The larger system size will lead to the peaks being less "defined" (as if averaging over 8 smaller-system tensile tests), but the general behavior is the same. Based on these results, a system size of 512 monomers is considered adequate to ensure that finite size effects are not significant.

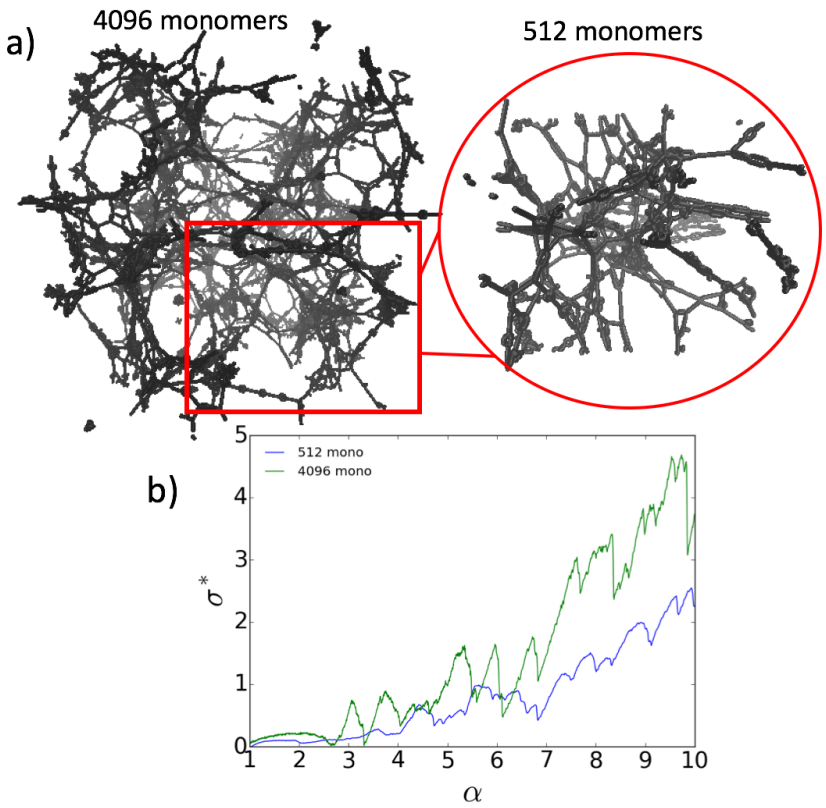

Figure 4.11: a) Connectivity of 4096-monomer and 512-monomer networks . b) Tensile response of both networks.

4.6.3.3 Calculating r_g for monomer "loops"

To calculate the average size of a free loop, we simulate a ring of beads whose length, n_{bead}, is equivalent to a loop formed by n monomers with arm length l. The appropriate "free" arms of the monomers are also simulated. When calculating the radius of gyration of the loop, r_g^{loop}, only beads along the loop "backbone" are considered. Figure 4.12 shows the calculated r_g^{loop} for a single loop in the simulation box for various n and l. It is found that $r_g^{\text{loop}} \propto n_{bead}^{0.67}$, which is the same scaling exponent as for free monomers. It is known that polymers rings, especially semiflexible rings, do not form spherically symmetric bodies [6,215]. Previous studies show that a contour length to persistence length ratio >10 ensures that a ring of semiflexible chains acts as a flexible ring, a condition satisfied by even our smallest rings (2-

loop, l_a=11). The asphericity [24], a measure of whether an aspherical shape is prolate or oblate, is ~0.67, indicating prolate spheroids, the shape taken on by flexible rings. Approximating the prolate spheroid as being roughly spherical (asphericity for these coils is ≈ 0.3), the volume of the coil will scale as $V_{coil} \propto n_{bead}^2$, as described in the main text.

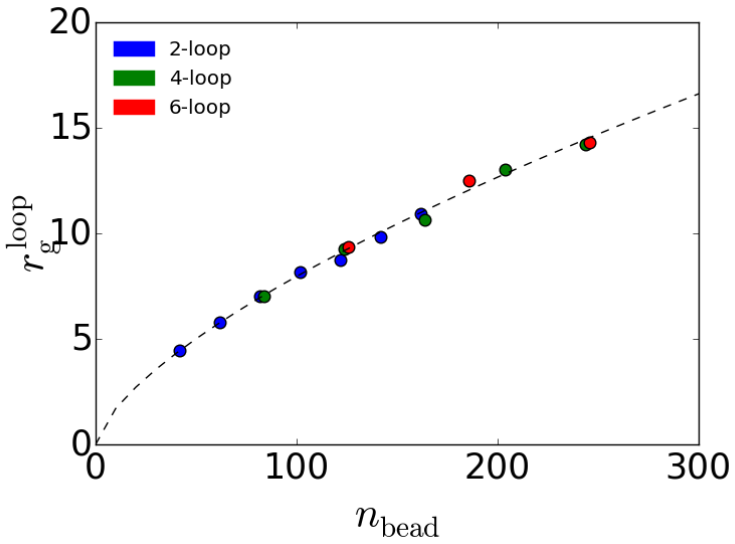

Figure 4.12: Plot of r_g^{loop} against n_{bead}. The dashed line corresponds to $r_g^{\mathrm{loop}} \propto n_{bead}^{0.67}$, the same scaling relation as the r_g for unreacted monomers.

5 Tuning the Sawtooth Tensile Response and Toughness of Multiblock Copolymer Diamond Networks

5.1 Introduction

Liquid crystalline elastomers (LCE), cross-linked networks composed of semi flexible polymer chains, orientationally align when uniaxially deformed at appropriate temperatures and densities (forming, e.g., nematic phases)[44,188], or when exposed to other stimuli such as electric fields[34], temperature[204], or light[124]. Translational alignment (e.g., smectic phases) occurs when chemically incompatible blocks make up the chains[63,234]. Anisotropic shape changes of the network occur when the chains orientationally order due to the prolate chain shape, a property central to artificial muscles[125,222] and high-strain actuators[218,232]. This macroscopic shape change highlights the coupling between material shape and chain orientation. For our ensuing discussion, the "director" in a LCE will refer to the direction in which the chains are preferentially aligned.

LCEs also display nonlinear elasticity, as seen in actin cytoskeleton networks[95,225], and spider dragline silk[110,228]. One reason for their nonlinear elasticity is the loss of entropic elasticity due to the existence of hairpins and other similar structures. For a chain hairpin, the entropy for unequal leg length is the same as that with equal leg length due to minimal transverse chain motion[3]. As such, the chain end-to-end distance does not correlate with the force, as only when hairpins unfold does entropy change.

LCEs can also undergo deformations at minimal or zero stress, a behavior referred to as soft elasticity or Goldstone modes. These soft deformations can be explained by the invariance in the free energy equations to changes in the director's direction[153]. Namely, during

deformation, cross-links can effectively “swap” places and leave the cross-link distribution, and entropy, unchanged. As such, the coupling of network shape to chain alignment causes zero force director rotations that result in zero force deformations.

Goldstone modes are but one of many ways how network connectivity can affect network behavior[18,19,28]. Two networks can show different tensile behavior just by changing the state in which they were cross-linked[211]. For example, simulations of actin networks show improved mechanical properties for networks with regular connectivity as compared to randomly cross-linked actin networks[38]. Realization of networks with regular connectivity has become more plausible as methods for "freezing" in mono domains[117] and cross-linking reactions that can be reversed[159], have reduced polydomains (i.e., regions with different directors that diminish mechanical properties[209]). Specifically, synthesis of diamond-like network architectures has been approached using tetra arm monomers (tetra-PEG monomers) that form a network by arm connection rather than through cross-linking[127]. Simulations confirm a more regular connectivity and experiments report improved mechanical properties using this method[175,197].

Simulations of LCEs with diamond connectivity and semiflexible block copolymer chains have shown a non-linear deformation behavior[4,17]. As hairpins in the network unfold during deformation, the system forms new segregated domains of cross-links to relax the stress and allow further extension. The breaking of modules has natural analogs in fibronectin[91,109], nacre[230], titin[171], and spider dragline silk[9,148]. Due to the repulsive interactions between block-A and block-B rich domains, there is an increase in interfacial energy when forming new domains. Accordingly, cross-links must form domains

concertedly to minimize the energetic penalty, giving rise to a very pronounced "saw-tooth" type of tensile behavior. Soft deformations manifest when the sample is deformed in a direction not aligned with the director. Previous work studied a symmetric ABA triblock chain (25%A-50%B-25%A) with a particular backbone stiffness. The number and height of stress peaks and the strain-orientation dependence observed in the tensile response of this case (summarized in Figure 5.1) is considered as a reference for comparisons with networks studied in the present work. Our main goal is to unveil basic principles for modulating the height and number of stress peaks on these LCE networks by using different backbone stiffness and chain architectures (i.e., the number and arrangement of blocks). Particular attention was given to networks that displayed behaviors that differed from that of the base network. The rest of the paper is organized as follows. Section 5.2 describes the simulation techniques and models adopted, including the methods used to obtain the properties of interest. Section 5.3 contains a discussion of the mechanical properties and deformation mechanisms for various chain designs (mainly ABA, ABCBA, and ABACA) subjected to uniaxial deformation. Finally, in section 5.4 several conclusions are presented, conjectured design principles for super-tough elastomers design are advanced, and areas for future explorations pointed out.

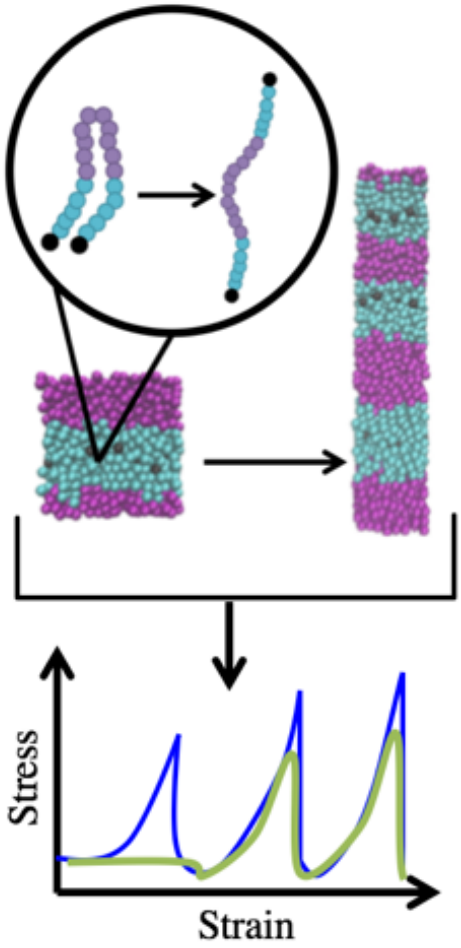

Figure 5.1: A broad overview of the previously studied system. Chains unfold and form new smectic domains, relaxing the stress and producing the saw-tooth shape to the stress-strain plot. The blue curve is for extension along the director while the green is for extensions not along the director.

5.2 Model & Methods

The network architecture consists of tetra-functional cross-links that lie on lattice points of a diamond unit cell when the network is fully expanded (see Figure 5.2). Each unit cell consists of 8 cross-links and 16 chains. Networks had either 2×2×2, or 4×4×4 unit cells, with the network bonded to itself across the periodic boundaries to simulate an infinite network. The chains are monodisperse 20-mers (unless otherwise stated), of coarse-grained Lennard-

Jones (LJ) type beads chained into blocks with different chemistry; Figure 5.3a shows examples of the main chain architectures explored. The diamond connectivity results in an entanglement free network, key to the mobility of the cross-links and the creation of the smectic domains.

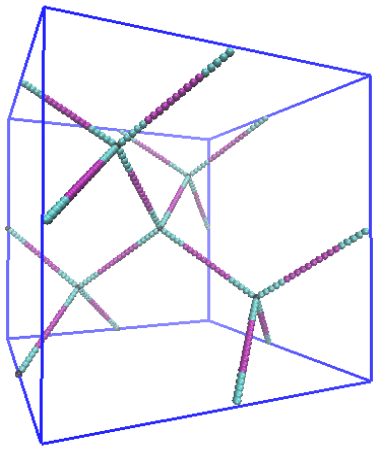

Figure 5.2: Fully swollen unit cell with 16, 20-mer chains and 8 cross-links in an ABA arrangement. The A, B, and cross-link beads are colored blue, purple, and black, respectively.

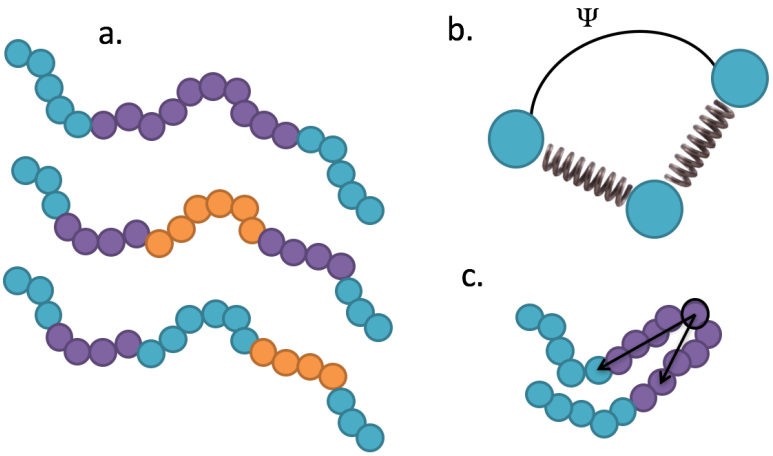

Figure 5.3: a) From top to bottom: ABA, ABCBA, and ABACA 20-mer chain architectures. The coloring scheme is as follows: A=blue, B=purple, C=orange, cross-links=black. b) The angle used in the bending potential is the angle between two adjacent bond vectors. c) Example of how the hairpin algorithm is applied. The bead circled in black is the bead being examined and the two arrows represent the vector between that bead and the two beads connected 4 beads away.

The specific potentials used in this work are chosen to represent a generic coarse-grained

polymer behavior, beginning with bonding that is modeled by the standard FENE potential:

$$U_{FENE} = -KR_0^2/2\ln\left(1-\left(\frac{r}{R_0}\right)^2\right)+4\varepsilon\left(\left(\frac{\sigma_{LJ}}{r}\right)^{12}-\left(\frac{\sigma_{LJ}}{r}\right)^{6}+\frac{1}{4}\right) \qquad (5.1)$$

Where the LJ-like potential term is cut off at $r=r_c=2^{1/6}\sigma_{LJ}$ to mimic excluded volume interactions and $K= 30\varepsilon/\sigma_{LJ}^2$ and $R_0 = 1.5\sigma_{LJ}$ are the spring force constant and maximum bond extension length, respectively. These widely used values of K and R_0 describe a stiff bond whose equilibrium length is approximately equal to the bead diameter ($\sim\sigma$);[71] such choices are not very influential on the properties of interest in our study. For non-bonded interactions, a cut and shifted LJ potential is used with the form:

$$U_{LJ} = 4\varepsilon^{ij}\left(\left(\frac{\sigma_{LJ}}{r}\right)^{12}-\left(\frac{\sigma_{LJ}}{r}\right)^{6}-\left(\frac{\sigma_{LJ}}{r_c}\right)^{12}+\left(\frac{\sigma_{LJ}}{r_c}\right)^{6}\right) \qquad (5.2)$$

for $r/\sigma_{LJ} \le r_c$ (U_{LJ} = 0 otherwise), where $r_c^{AA} = r_c^{BB} = r_c^{CC} = 2.5\sigma$ (attractive potential), and $r_c^{AB} = r_c^{AC} = r_c^{BC} = 2^{1/6}\sigma$ (repulsive potential) for triblock and pentablock copolymers (A blocks and cross-links are taken to be the same type), and $\varepsilon^{ij}=0.5\varepsilon$ for all interaction pairs ij. These choices of r_c cutoffs and ε^{ij} values are chosen so that they result in an effective Flory-Huggins parameter of $\chi N \approx 32$.[83]

Mesogenic behavior is induced by a bending potential that is applied to all beads that are not directly bonded to a cross-link, which has the form:

$$U_{angle} = k_b T K_{bend}\left(1+\cos(\psi)\right) \qquad (5.3)$$

where K_{bend} is a stiffness constant and ψ is the angle between two consecutive bond vectors. To work in reduced LJ units, reported properties are scaled as follows: $T^*=Tk_b/\varepsilon$, $\tau=t(\varepsilon/(m\sigma_{LJ}^2))^{1/2}$, $\sigma^*=\sigma\sigma_{LJ}^3/\varepsilon$, $P^*=P\sigma_{LJ}^3/\varepsilon$, $\gamma=\Gamma\sigma_{LJ}^3/\varepsilon$, and ρ = (# of particles)/ σ_{LJ}^3 , where T is

temperature, σ^* is stress, P is pressure, Γ is the toughness of the material, ρ is number density, and all scaling variables (σ_{LJ}, m, ε) are set to 1. For all our simulations we use a constant value of T*= 2.0 which is taken to correspond to room temperature and is a choice which will set the chain length of a given polymer needed to match a prescribed χN of the model. For reference, we cite specific polymer chemistries that approximately map select features of our coarse grained model. Poly(2,5-di(2′-ethylhexyloxy)-1,4-phenylenevinylene) (DEH-PPV), is a semiflexible polymer that when linked with a flexible block of 4-vinylpyridine (4VP) gives a block copolymer that has a $\chi N \approx 15$ at room temperature (by estimating the order-disorder transition temperature as $\chi N \approx 10$). [119] Coarse graining the DEH-PPV and 4VP such that each bead has the same size ($\sigma_{LJ} \approx 12$ Å) leads to ≈2 DEH-PPV monomers per coarse-grained bead in a 20-mer ABA chain. This also sets a bead mass of *m* ≈ 716 Da, a parameter that only influences dynamic quantities, not equilibrium properties. Since DEH-PPV has a persistence length of ~110Å,[177] then $K_{bend} \sim 5\text{-}10$ as estimated from: persistence length $\approx K_{bend}\sigma_{LJ}$, a relation appropriate for chains of hard tangent beads.[55] (Note that the stiffness of a given polymer backbone can be increased by larger or bulkier grafted sidechains[23]). This mapping of χN and all bonded and non-bonded parameters given above (i.e., K_{bend}, σ_{LJ}, etc.) is non-unique since they depend on the level of coarse graining and what real temperature is chosen to represent T*=2.0.

We emphasize that we purposely chose a generic polymer model whose potential functions and parameter values are based on well-established models that have been extensively used and shown to reproduce the static and dynamic behavior of many polymer systems.[70] Such a well-known class of coarse-grained models is appropriate not only because it is well-suited to capture physical trends of polymer behavior, but also because it allows us to build on the

existing body of work by studying systems with added features of molecular complexity. Indeed, we build on previous studies that have used similar models to describe flexible homopolymers, block copolymers,[83] athermal rigid chains,[55] and networks. We also note that "reverse" mapping of a coarse-grained model into a specific chemistry is non-trivial since, as an underdetermined problem, many different polymer chemistries (including purposely designed ones) can potentially be mapped that way (and with varying degree of fidelity). However, if one were interested in targeting a specific chemical system, there exist several well-known strategies for developing a coarse grained model from atomistic models,[30,194] including some that use similar potentials as those used here.

All simulations were performed in a strain driven framework, beginning with a fully swollen lattice that is compressed to a melt density, and equilibrate using two different techniques. ρ=.9 is chosen as the melt density as similar densities are used in other particle-based polymer simulations.[30,83] For 2×2×2 unit cells the system is compressed directly to the melt density at a constant temperature T*=2 in 10^6 time steps. For 4×4×4 unit cells the system is compressed in two stages: a compression to a density of ρ=0.5 in 10^7 time steps, followed by a slow compression to ρ=0.9 in 5×10^7 time steps. The compression follows $l=l_0\exp(k\tau)$, where l is the simulation box length, l_0 is the original box length, k is a dimensionless compression rate, and τ is the scaled LJ time. The slow compression from ρ=0.5 to 0.9 in the 4×4×4 unit cell simulations allows for faster equilibration as chains can rearrange more easily. The dynamics of the 2×2×2 unit cell simulations were fast enough that the slow compression was not needed. A variation of these equilibration protocols was used when complete morphological equilibration was suspect: rather than enacting the stipulated LJ potentials for the non-bonded interactions from the outset, all such interactions

were first set to be purely repulsive, WCA interactions as the system was compressed. Subsequently, the temperature was raised to T*=3 and the stipulated LJ interactions enabled, after which the system was cooled from T*=3 to T*=2 linearly over a cooling time of $\sim 10^7$ time steps. In all simulations, the temperature is controlled using a Nosé-Hoover thermostat, and the equation of motions are integrated using a velocity Verlet algorithim, with a timestep equal to .005τ.

For all methods described above, once the protocol is completed, the network is allowed to equilibrate until we see steady state values for the bending, non-bonded, and bonded potentials, as well as the pressure tensor components, p_{xx}, p_{yy}, and p_{zz}. The time scale of equilibration is $\sim 10^8$ time steps. This state is considered the undeformed state, α=1, where α is the extension ratio, i.e., the length of the longest box edge scaled by the undeformed cubic box length (l_{ext}/l_0). To simulate uniaxial deformations, the box length along the deformation axis is expanded linearly every 5000 time steps up to α=5 in 10^8 time steps, leading to a deformation rate of 8×10^{-6} α/τ, while the other two dimensions are contracted to keep a square cross section and the volume constant. The stress is calculated using:

$$\sigma^* = -P_{ii}^* + \lambda\left(P_{jj}^* + P_{kk}^*\right) \quad (5.4)$$

Where P_{ii}, P_{jj} and P_{kk}, and λ are the extensional direction pressure tensor component, the non-extensional direction pressure tensor components, and the Poisson ratio (λ=0.5), respectively. To quantify the chain alignment in the system, we calculate the bond order parameter P_2, which is defined as the largest eigenvalue of the matrix:

$$Q_{ij} = \frac{1}{N}\sum_{n=1}^{N}\frac{3}{2}\hat{u}_{in}\hat{u}_{jn} - \frac{\delta_{ij}}{2} \quad (5.5)$$

Where $\hat{u}_{in}$ is unit vector pointing along bond *n* in the *i* direction, *N* is the number of bonds in the system, and δ_{ij} is the kronecker delta.

The toughness or the energy absorbed during deformation, is calculated using:

$$\gamma = \int_1^{\alpha_f} \sigma^* H\left(\sigma^*\right) d\alpha \qquad (5.6)$$

where $H(\sigma^*)$ is the Heaviside function, which is introduced to ignore negative stresses. We ignore the negative stresses because we are interested in the total energy absorbed (i.e. total work done on the system), and the energy "returned" (work done by the system) when stress is negative would essentially be dissipated. Conventionally, the upper limit of integration in Eq. 5.6 should be α at fracture; however, since our model bonds do not account for breaking events, α=5 was chosen for convenience. To provide some calibration for the toughness values to be reported, we can compare the simulated toughness of ~0.1GPa[217] (calculated for extension from α=1 to 5) for an atomistic model of a polyethylene network with diamond connectivity to γ=0.36 for a fully flexible coarse-grained homopolymer network of comparable molecular weight; hence a unit in γ in our model would correspond to ~0.28GPa.

The number of hairpins in a network during deformation is calculated by first drawing two vectors from a central bead in a chain to a bead n_{vector} bonds away in each direction along the chain (see Figure 5.3c). If the angle between the two drawn vectors is less than a threshold angle, the chain in question is labeled a hairpin. We apply this algorithm over all possible beads on all chains, using n_{vector}= 4 and a threshold angle of 60°. Note that the exact choice of these parameters is arbitrary and that the algorithm does not discern between a chain having one or two hairpins, although the occurrence of double hairpin chains is negligible. For ABA networks, if the chain has an angle within the threshold angle or both chain ends reside in the same segregated domain of cross-links, then the chain is labeled as a hairpin.

Repeated runs of the same network with differing starting configurations give slightly

differing results. As such, we show representative plots and report the average values where possible. Stress-strain, hairpin, and order parameter plots show values averaged over ~0.005-0.01α.

5.3 Results & Discussion

Our results are organized into three groups of networks with constituent chains of increasing complexity: (i) ABA chains of varying composition and fixed stiffness, (ii) ABA chains with varying stiffness, and (iii) pentablock chains with three different chemical blocks.

5.3.1 Compositional scan with constant stiffness

5.3.1.1 Stiff chains tend to form lamellar morphology

For symmetric ABA chains (Figure 5.3a, top chain), the main variable is the composition, the volume fraction of A block to B block. Tests were performed on 2×2×2 unit cell networks, varying f from 80%A to 10%A in increments of 10%. K_{bend} and chain length were kept constant at 4.0 and 20 beads, respectively. A lamellar morphology was observed for all networks throughout the entire deformation, regardless of composition, which is surprising as block copolymers are known to form curved interface morphologies (cylindrical, spherical, etc.) as block asymmetry is increased[111]. In our systems, however, those morphologies may be precluded in part because the chains are cross-linked at their ends, and partly because curved interfaces would incur a chain bending energy penalty, offsetting any reduction in interfacial energy. In the supplemental information we present results from simulations of uncrosslinked chains showing that for 20%A the stable morphology is not a lamellar phase as in the crosslinked network but rather: (i) a weakly segregated, perforated lamellae if K_{bend}=4, and (ii) a disordered phase if K_{bend}=0, in agreement with previous

studies.[83]

The tensile behavior of all these networks was qualitatively consistent with that of the base network (i.e., Figure 5.1). Figure 5.4 shows stress-strain curves of extension in the direction normal to the lamellae at α=1 for several compositions. The stress at α=1 is often not exactly 0 because of residual stresses that anisotropic morphologies develop inside a cubic simulation cell. We also note that the stress plots shown are for individual deformation runs and hence significant sampling variance when runs are repeated with different initial configurations. The trend in the tensile response as a function of composition is treated in section 5.3.1.2.

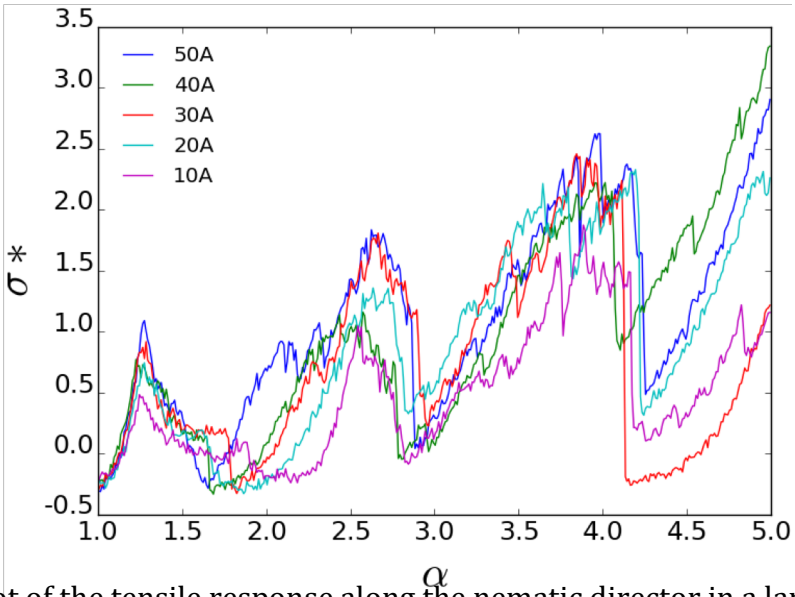

Figure 5.4: Plot of the tensile response along the nematic director in a lamellar phase from ABA networks of different compositions (from 50%A to 10%A).

Figure 5.5 plots the stress-strain response for an extension orthogonal to the nematic direction for the reference system. Before hardening there is a dip in the stress, known as buckling. The free energy equations of these materials[219] (extended from classical rubber theory) predict a drop in energy before hardening. Buckling would normally manifest as

spontaneous expansion in the non-extended directions, however, it manifests as a drop in the tensile force because here the total volume is kept fixed. Buckling corresponds to realignment along the extensional direction of both chain orientations and the lamellae plane normals. Reorientation entails stretching of the lamellae, followed by lamellar splitting into equal parts, and finally rearrangement into new, reoriented lamellae, as depicted in Figure 5.5.

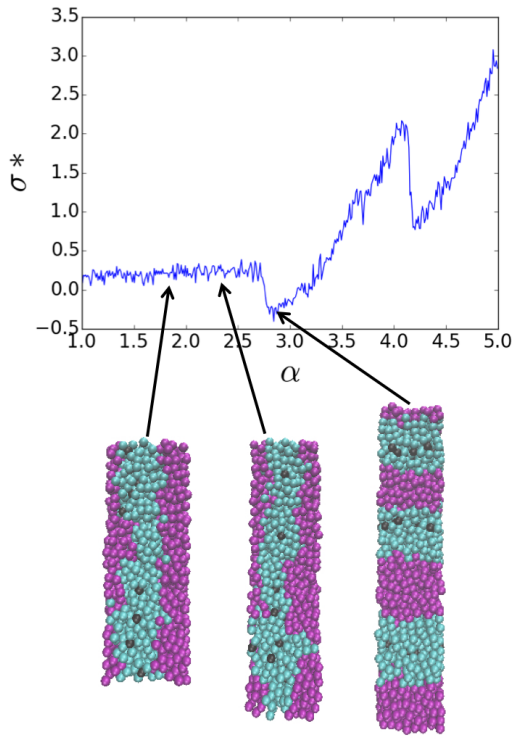

Figure 5.5: Tensile response for extension not along the initial director with sample snapshots at selected deformations. See Figure 5.3 for color scheme.

5.3.1.2 Effect of Composition on Reference Network Toughness

The various networks show a maximum in toughness as a function of composition. We observe a maximum in γ at ~50%A suggesting that A-B interactions, maximal for the 50-50% composition, may play a key role. A simple model is formulated below to explain our observations. Consider the changes occurring over a range of α where a single saw-tooth

occurs as depicted in Figure 5.6, where a two-domain lamella is transformed into a four-domain structure. Assume that energy changes associated with such a transition primarily arise from changes in LJ interactions, while contributions from bending and bonding interactions are negligible, as described previously.

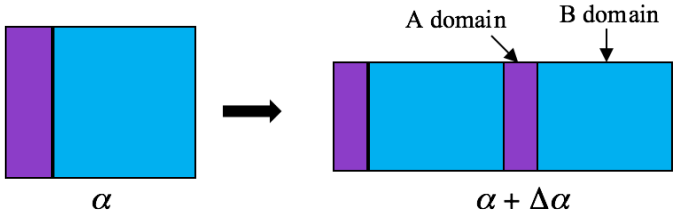

Figure 5.6: Depiction an idealized strain-induced transition from 2-domain to 4-domain lamellae structure for network made of symmetric ABA chains.

If we denote n_A (n_B) as the number of A (B) beads that came in contact with B (A) type beads during such a transition, then the energy change per unit volume associated with this $\Delta\alpha$ will be:

$$\frac{\Delta U_{tooth}}{V} \approx \frac{1}{V}\sum_{i=1}^{n_A}\sum_{j=1}^{n_B} u_{ij} \quad (5.7)$$

where u_{ij} can be taken as being proportional to the χ parameter; i.e., the energy change associated when A-A and B-B (bead-bead) contacts are exchanged by A-B contacts:

$$u_{ij} = u_{AB} - 1/2\left(u_{AA} + u_{BB}\right) = kT\chi \quad (5.8)$$

Hence, Eq. 5.7 becomes:

$$\frac{\Delta U_{tooth}}{V} \approx \frac{kT\chi}{V} n_A n_B \quad (5.9)$$

If we now assume that n_A and n_B are proportional to the relative number of beads in their respective domains, namely, $n_A \sim f_A$ and $n_B \sim (1\text{-}f_A)$, and that the number n_{teeth} of such

transitions is the same for a given total α (regardless of f_A), then the total toughness $\gamma = n_{teeth} \times \Delta U_{tooth}/V$ can be expressed as:

$$\gamma \propto \chi f_A(1 - f_A) \qquad (5.10)$$

This model assumes that the tensile response is solely determined by inter-block interactions and hence by χN, while neglecting chain bending. For comparison to simulation data, Figure 5.7 plots the calculated toughness for deformations in the nematic direction, normalized by the γ for the 50%A case, as a function of composition. In calculating average toughness values, we omit responses for extensions orthogonal to the nematic direction to reduce the variability associated with the frustrated buckling. Equation 5.10 predicts a maximum toughness at 50%A and an overall trend that shows reasonable qualitative agreement with simulation data. Although chain stiffness also influences χN and can lead to entropic segregation, as a first approximation our model neglects the effect of chain stiffness relative to enthalpic contributions. This can be justified by the fact that the toughness of athermal homopolymer networks (experiencing purely entropic contributions) with K_{bend}=4.0 is ~0.4, which is significantly smaller than that of the reference network (γ~4.8). If the simulation data in Figure 5.7 were adjusted to remove the entropic contribution, the fit would improve. The study of how entropic segregation affects the tensile response lies beyond the scope of this work and will be the topic of future research.

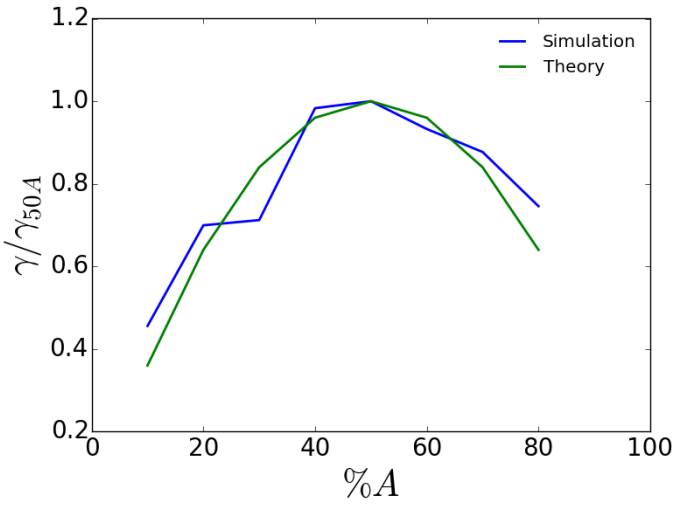

Figure 5.7: Plot of γ as a function of composition (%A) comparing simulation to the predictions of Eq. 5.10. All values are normalized to $\gamma_{50\%A}$.

5.3.1.3 The Effect of T and Deformation Rate on γ*

The tensile response should also be a function of both temperature (on account of its effect of entropy on elasticity and that χN∝1/T), and the strain rate (which sets a deformation time scale that competes with the microstructural relaxation time scale). In the supplemental information (Figure 5.19), it is shown that for the range of strain rate and temperature tested, the stress response is not systematically affected by the strain rate, while the toughness tends to decrease with temperature, likely through its effect ont χN (i.e., a reduction of interfacial energy).

5.3.2 Lower Stiffness Networks

5.3.2.1 Morphologies With Curvature Isotropize the Orientation-Dependent Tensile Response

Due to the strong bias of K_{bend}=4 networks to form lamellar phases, tests similar to those in section 5.3.1.1 were performed with K_{bend}=2.5. The most symmetric compositions, 50%A and 40%A, still show lamellar morphology in the undeformed state. The 30%A system forms instead a perforated lamellae, hereafter to be referred to as the "perforated network", with

the lamellae normal pointing in the [110] direction. This perforated network is seen consistently using different preparation protocols with 2×2×2 unit cell systems. For a 4×4×4 unit cell network, two different types of partially disordered morphologies were found in the undeformed state (see Figure 5.8). One is perforated lamellae with lamella normal pointing in the [1 -1 2] direction, consistent with the perforated network from the smaller system. The other is a lamellae-like morphology with connections between the lamellae, which we refer to as the "defective" lamellae network. The difference in total energy between these two morphologies is small, ~0.0145kT per LJ bead, which precludes singling one of them out as being more stable (on account of possible differences in entropy).

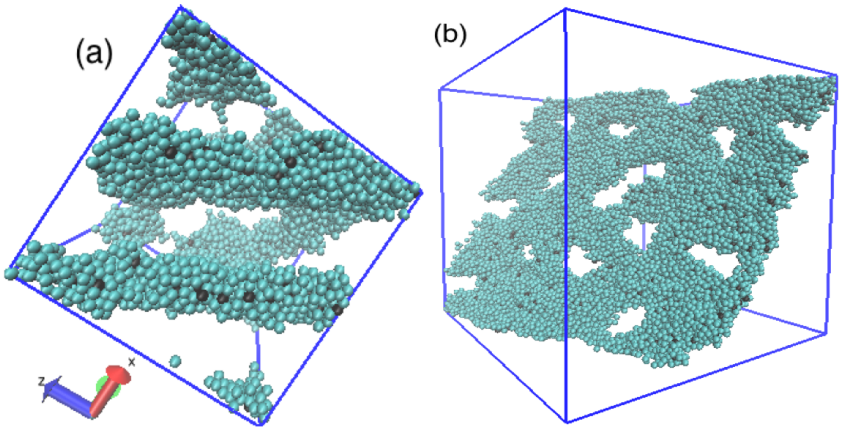

Figure 5.8: Morphologies obtained for 30%A, K_{bend}=2.5 system with a larger lattice. For clarity B beads are removed; see Figure 5.3 for color scheme. (a) is the defective lamellae and (b) is the perforated lamellae pointing in the [1 -1 2] direction. (a) is rotated to show the lamellar morphology and how they are interconnected. (b) is a depiction where the fragments of the lamellae in the periodic box are "stitched" together to form one single lamellae. We use extra periodic images to create a more complete lamellae. B block is removed for clarity, see Figure 5.3 for color scheme.

Despite their differences, the tensile responses for both initial morphologies shown in Figure 5.8 are consistent with each other (see example in Figure 5.9a), showing isotropization of the tensile response as compared to the base network; i.e., the stress response becomes

largely independent on the strain direction. Lack of one universal director, reflected by the low values of P_2 (Figure 5.9b) in the undeformed state, indicates that chain orientations are not well correlated. The distribution of dot products between chain bond vectors and the calculated director in the underformed state (Figure 5.9c) shows a much broader distribution for the perforated network, populating low alignment states that are not seen for the reference network. Thus, soft deformation in perforated networks will always appear (regardless of extension direction), which differs from soft deformation in lamellar morphologies where there is a universal director, as seen in Figure 5.5. After buckling, the perforated lamellae become untilted and unperforated, with the subsequent sawtooth tensile response being consistent with that of the base network. Isotropization is a product of both the realignment of chains along the strain direction and the morphology switching to lamellae. Values of γ are 1.43 and 1.34 for the perforated and defective lamellar networks, respectively; roughly a third of the γ for the base network. This reduction in toughness is a result of a less symmetric composition (as per Eq. 5.10), a lower K_{bend}, and the absence of one stress peak due to the initial soft response.

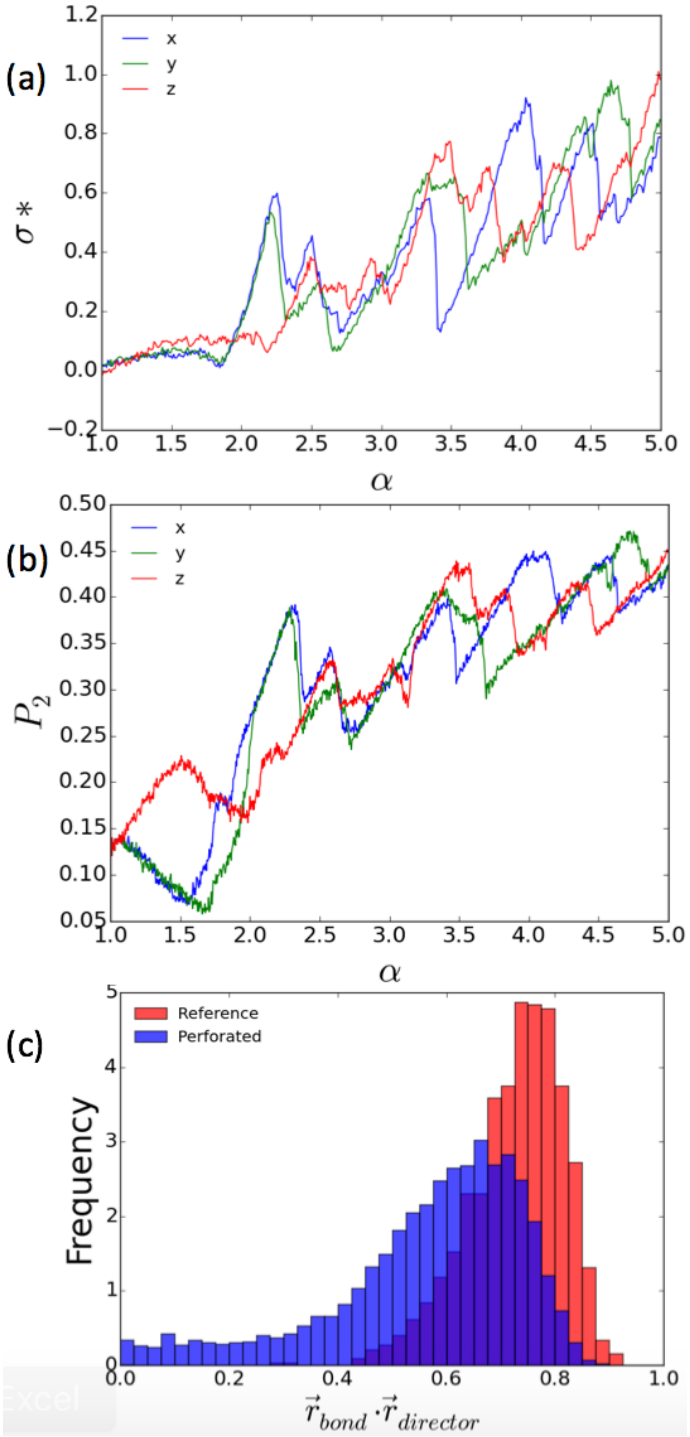

Figure 5.9: Plots for tensile response (a), P_2 (b), and chain alignment to the A-B interface (c). normal (c) for the perforated network. The distributions in (c) are averaged over 20 snapshots spanning 2×10^6 timesteps.

The soft behavior for the 4×4×4 unit cell perforated networks also differs from that of the

reference network. As outlined in Figure 5.10, lamellae first become stretched and split along perforations that reform into a structure similar to the defective lamellar morphology. Further extension rearranges the network into untilted, unperforated lamellae. The defective lamellar morphology rearranges directly into untilted, unperforated lamellae.

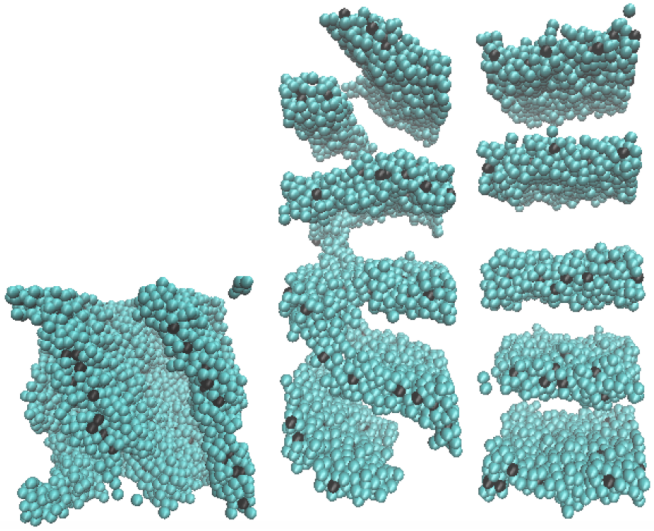

Figure 5.10: Soft response of the perforated network for α = 1, 2, and 2.25 (from right to left). B block removed for clarity; see Figure 5.3 for color scheme.

Simulation results (see SI) for smaller perforated networks confirm the above results.

A network with 30%A, ABA chain structure, K_{bend}=5.0 and chain length of 40 beads was also simulated to further confirm the prediction of a perforated network since this new network should be similar to the one described above for having the same ratio of persistence length to contour chain length.[48] The undeformed morphology observed for this network is untilted, perforated lamellae which upon deformation also exhibits stress-response isotropization (see Figure 5.20 in the SI), indicating that such a response is independent of the tilting of the lamellae. All these results suggest that any morphology whose geometry leads to weak director alignment, will show some degree of isotropization. Morphologies with interfacial curvature do not have a uniform director, a feature only present in the

lamellar morphology. As a consequence, the exact undeformed configuration, even if partially disordered, may not be critical for determining the tensile behavior of these systems, since, after the initial soft stress response, the system will "forget" its initial state.

5.3.3 Pentablocks

A pentablock chain only has 5 unique architectures in this model; ABABA, ABCBA, ABACA, ABACB, and ABABC. We only discuss ABCBA and ABACA. ABABC and ABACB chains are omitted because their behavior is not saw-tooth-like as the end blocks are not the same type. This disparity disrupts the ability for a cross-link to move into a new layer and maintain phase segregation. ABABA chains act as extensions of ABA chains and are omitted.

5.3.3.1 ABCBA: stronger tensile response through increased number of interfaces

For all stiffness parameter values and compositions tested for an ABCBA chain (30%A-40%B-kb4, 20%A-70%B-kb4, and 20%A-70%B-kb2.5), the lamellar morphology was observed in the undeformed state (as depicted in Figure 5.11). Because the chain has central symmetry, the mechanism of deformation is just an extension of that of ABA chains. The only deviations occurred during the formation of new domains, where the time scale of morphological rearrangement was now too slow to keep up with the strain rate, an effect of the short simulation time scales, leading to some blunter peaks. Compared to the base network, there is an approximate 30% increase in γ for a 30%A-40%B, K_{bend}=4.0, 20-mer chain network, a consequence of a larger number of interfaces that enhance the resistance to deformation. As the volume fraction of one of the blocks approached 0, the effect of the added interfaces diminished and γ decreased to values similar to those for ABA chains. These results suggest that any multiblock chain with center symmetry will act as an ABA chain with a suitably increased number of interfaces.

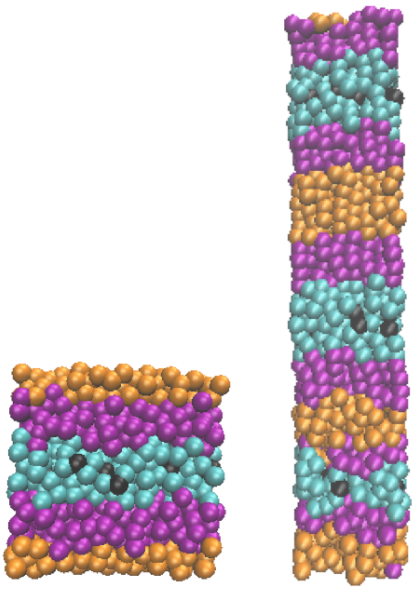

Figure 5.11: Morphology of 30%A-40%B-30%C ABCBA network for the undeformed (left) and α=3 (right) states. For the exact chain topology of this network and color scheme see Figure 5.3a, middle chain.

5.3.3.2 ABACA: multiple folding points lead to increased transformation rates

The most unique new behavior among pentablock chains is observed for ABACA chains (chain topology shown as the middle chain in Figure 5.3a). The network consists of 60%A, 20%B , 20%C, and 20-mer chains with K_{bend}=4.0 (referred herein as the ABACA network). We arrange the chains so that all 4 chains connected to one cross-link are oriented as either ABACA or ACABA with respect to the cross-link (see Figure 5.12). Due to the connectivity of the network, by setting one cross-link to have a certain orientation of the chains, this automatically sets the chain orientations of all other cross-links. Due to the lack of chain central symmetry, the well equilibrated undeformed morphology tends towards an interconnected lamellar morphology (Figure 5.13), akin to the defective lamellae network discussed before (see Figure 5.8). The interconnections allow hairpins to form and some chains to escape confinement in a lamellae domain whose spacing (14.2) is shorter than the

chain contour length (19.2). In our ABACA network morphology, the blocks arrange in a ABACABA sequence, which can be compared with the ...ABAC... equilibrium arrangement observed for experimentally synthesized ABACA type block copolymer (uncrosslinked);[61] this small difference can be ascribed to the topological constraints introduced by the crosslinks in the networks.

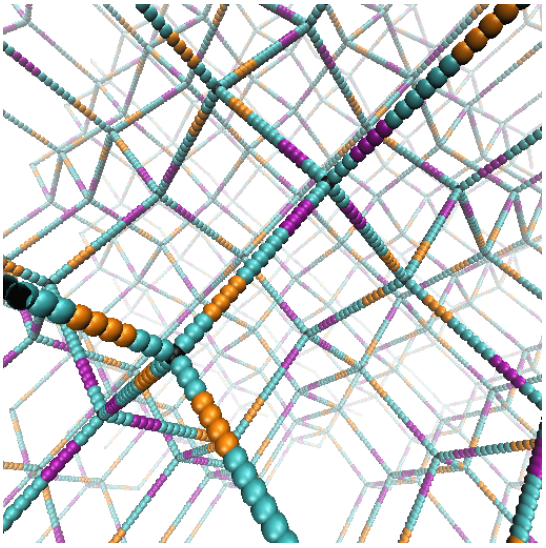

Figure 5.12: Snapshot of fully swollen ABACA network, showing the chain orientation flip between neighboring cross-links. See Figure 5.3 for color scheme.

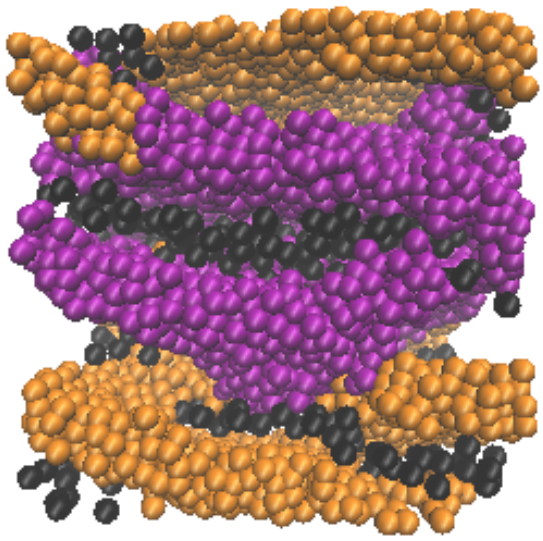

Figure 5.13: Snapshot of the undeformed ABACA network. The morphology shows interconnections similar to the perforated network. Block A is omitted for clarity; see Figure 5.3 for color scheme.

The chain asymmetry is also responsible for an increase in the number of segregated domains (as seen in Figure 5.11) and stress peaks (as observed in Figure 5.14) for a given strain. Indeed, chain asymmetry leads to shorter hairpin lengths (the length of the shortest hairpin leg) because the chains fold across either the B or C block, and not across the chain center. When the hairpin unfolds, these shorter hairpins reduce the change in network length, and thus, reduce the relaxation caused by new domain formation. As a result, stress rebuilds within a smaller range of α, and a larger number of transformations occurs over a given deformation range. The value of γ ranges from 5.70 to 12.80 for 2×2×2 unit cell systems and from 4.90 to 6.25 for 4×4×4 unit cell systems, with averages of 8.48, and 5.59, respectively. The smaller system's toughness is approximately twice that of the reference network, a consequence of two factors: The increase in the number of peaks, and a larger contribution from stress upturns. The larger number of peaks allows the system to create more segregated domains of cross-links (8 at α=5), and maintain a greater on-average stress.

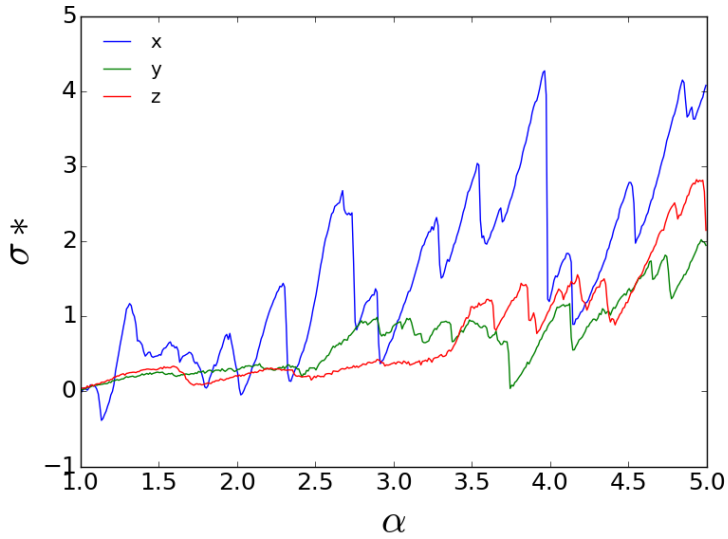

Figure 5.14: Tensile response for the ABACA-60A20C-kb4-20 network. An increase in number of teeth is seen in all directions.

The upturns occur when cross-links are pulled across very long distances, incurring in large energetic penalties, and thus, large stresses. Despite the higher toughness, smaller systems show larger variability than larger systems, likely due to finite size effects.

Chains can have three main types of conformations during deformation: (i) U-shaped, with both ends in the same segregated layer, (ii) unfolded, or (iii) a hairpin. (i) and (iii) are equivalent in the base ABA network but not so in the ABACA network due to chain asymmetry and the doubling in the number of segregated domains, where hairpins chain ends can end up in two distinct cross-link layers as depicted in Figure 5.15. As described below, the existence of additional or secondary cross-link domains has a unique effect on the folding/unfolding of hairpins in the network during deformation.

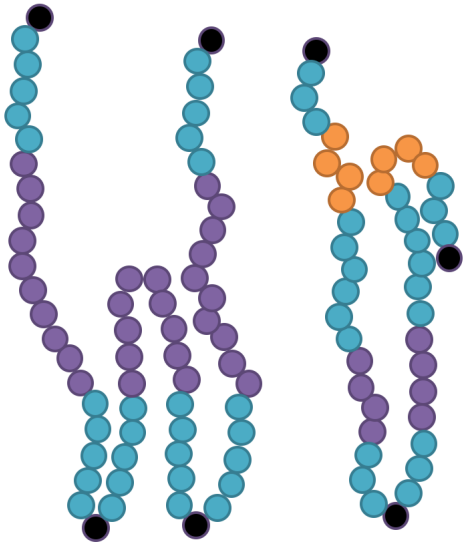

Figure 5.15: Cartoons of the chains in a deforming network for the reference case (left) and the ABACA system (right). The ABACA chain has a shorter hairpin which leads to the formation of secondary sets of cross-links. See Figure 5.3 for color scheme.

5.3.3.3 Shortened hairpin folding length leads to restoration of the hairpin population

To probe the microscopic behavior of chains, the fraction of hairpins η with respect to the total number of chains, was calculated throughout the deformation. Figure 5.16a plots the evolution of η for the reference network, showing that it follows the expected microscopic picture: the initial number of hairpins progressively drops as the system forms new layers. Figure 5.16b shows that the ABACA network follows a very different and unexpected trend where the number of hairpins can either stay constant, or even increase during domain formation. To understand this behavior, direct visualization was used to identify the underlying microscopic events.

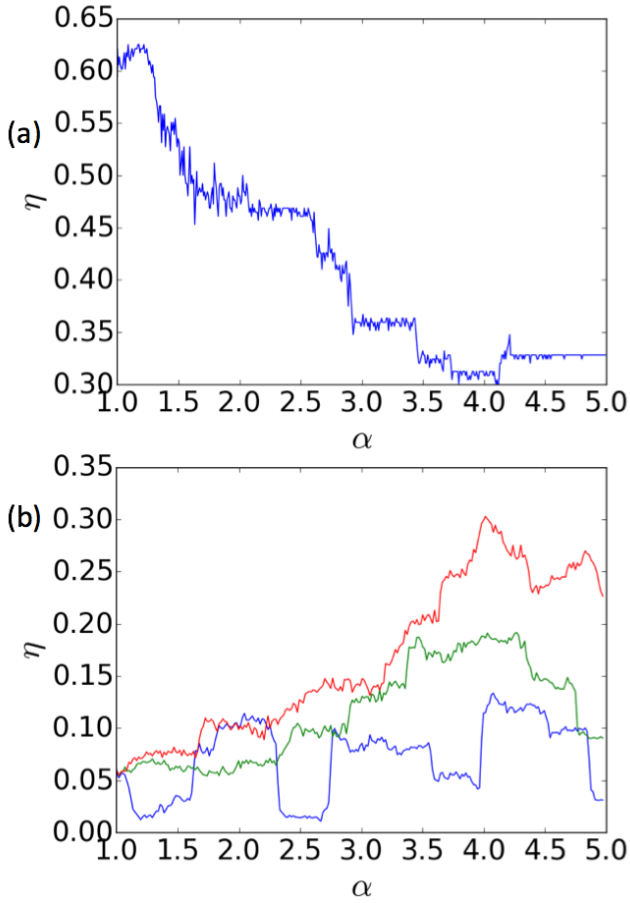

Figure 5.16: Fraction of hairpins as a function of α for: a) the base network, and b) the ABACA network.

Figure 5.17 depicts a center cross-link with the 4 chains bonded to it, at increasing values of α. The cross-link begins with one hairpin becoming unfolded as α increases (states 1 to 2), while movement of other cross-links pulls on the blue and pink chains, causing relaxation by folding of the black and purple chains (states 2 to 3, a "restorative" event). A large displacement of the central cross-link unfolds the black and purple chains and folds the blue and pink chains (states 3 to 4, a "transfer" event). These processes account for the unique trends in hairpin population seen in Figure 5.16b, and result from the secondary set of cross-links depicted in Figure 5.15 which effectively reduce the distance between domains of cross-links compared to the base network. This restorative behavior is not exclusive to the

ABACA system as a small number of these events can also be detected in the base network, but the reduced cross-link layer spacing makes such events more prevalent in ABACA networks.

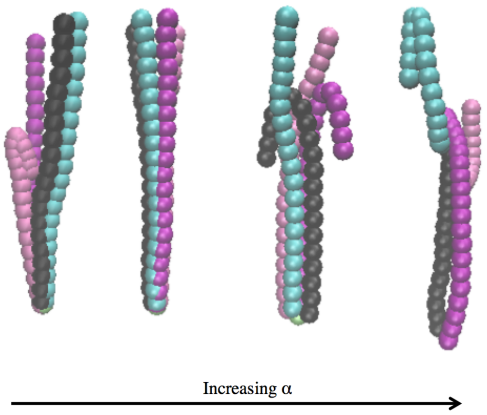

Figure 5.17: A progression of 4 ABACA chains connected to the same cross-link going from α= 1.5 (state 1), to 2.2 (state 2), 2.5 (state 3), and 3.0 (state 4).

The tensile response of the ABACA network also differs from that of the reference network in other respects. First, new smectic domains formed during deformation are not made up of purely B or C blocks. This occurs because the asymmetry of the chains coupled with the additional sets of secondary cross-link layers make it difficult to achieve complete microphase segregation. Second, these networks also exhibit peculiar soft behavior consisting of intermediate lamellar states. Indeed, as lamellae are stretched, they rearrange into structures that are tilted with respect to the extensional direction and that repeatedly rearrange into less tilted structures until the perpendicular lamellae are formed (see Figure 5.18). These tilted structures originate from dimensional mismatch between the simulation box and the lamellae unit cell. As more lamellae are formed, they become thinner, decreasing the unit cell dimension, which reduces the tilt. We see evidence of this effect in other systems

but it is magnified in this system as lamellae include both B and C layers. All these peculiarities stem from the chain asymmetry and how they fold and unfold

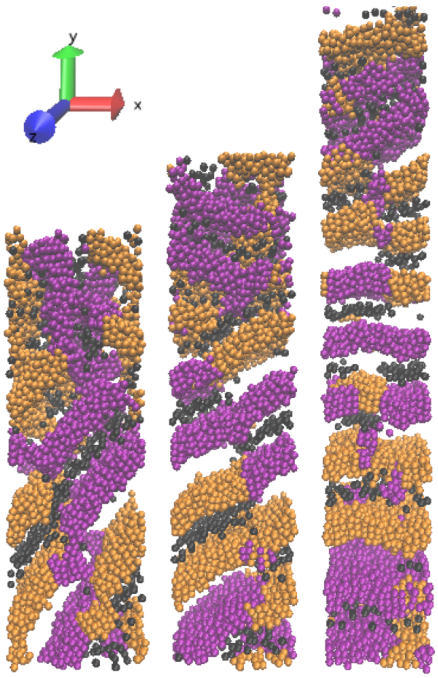

Figure 5.18: ABACA-60A20C-kb4-20 network undergoing soft deformation. From left to right the values of α are 2.2, 2.6, and 3.5, respectively. The lamellar tilt decreases as the deformation proceeds. Block A is removed for clarity; see Figure 5.3 for color scheme.

5.4 Conclusions

We extended previous work on LCE with diamond connectivity whose tensile response exhibited a characteristic sawtooth shape and an anisotropy with respect to the strain orientation, to identify two network designs whose tensile behavior has distinct features from that of the original, base network. First, a symmetric, triblock ABA chain with 30%A volume fraction and lower chain stiffness (k_{bend}=2.5) gives rise to a perforated lamellae phase in the undeformed state. Compared to the base network (whose ABA, 50%A, k_{bend}=4 chains form a lamellar phase in the undeformed state), this "perforated network" lacks a common director, requiring local director realignment along whichever direction the

network is being deformed. After buckling, the tensile responses become independent of strain direction. It is this combination of the domain reorientation and lamellar morphology after buckling that leads to isotropization of the stress response to the strain direction. It is conjectured that any morphology that (in the undeformed state) lacks a universal director should show some degree of isotropization of the tensile behavior. As an example, two very different initial morphologies for the perforated network were found to exhibit very similar tensile responses. Further studies could target networks whose AB morphologies have 3-D symmetry (like spheres or bicontinuous phases) and see whether their deformation behaviors approach those reported here; e.g., regarding isotropization and early buckling. Preliminary tests of networks with different stiffness values for each block have produced a cylindrical phase which also displays isotropization. Bicontinuous networks would be difficult to achieve due to the high curvature and the need to exactly match the simulation box to the unit cell of the morphology[129,130].

Second, a nonsymmetric, pentablock ABACA chain with 60%A volume fraction, 20%B volume fraction, and k_{bend}=4 gives rise to networks with several unique features. Compared to the base network, these ABACA networks exhibit an increased number of smectic domains for a given strain, a consequence of shorter, asymmetric blocks, which leads to both smaller cross-link domain spacing and shorter hairpins. Shorter hairpins and domain spacing couple to reduce the network deformation required to induce the formation of new domains. Another effect of the decreased hairpin length is the formation of secondary cross-link domains (Figure 5.15), which promotes abundant restorative and transfer hairpin events that counter the decrease in number of hairpins as α increases (Figure 5.16a). It would also be interesting to investigate whether a chain can be designed to give three hairpins, as a way

of inducing more transformations, which could lead to a tougher network.

Beyond the specific chain architectures and behaviors described above, some general design principles can also be put forth. First, the tensile response of multiblock chains with central symmetry will be an extension of the base network. Center symmetric chain networks are further toughened by increases in the number of interfaces (e.g., ABCBA chains) that must be broken when forming new domains, with toughness being maximized when the composition is as close to symmetric as possible. Second, isotropization of the stress response can be achieved by favoring undeformed morphologies lacking a common director. Non-uniform directors are associated with curved interfaces (appearing at lower chain backbone stiffness) as curvature precludes the ability of chains to align in one direction. Third, chain asymmetry, coupled with shorter blocks, can increase the number of teeth in the saw-tooth response per unit deformation.

Some of our ongoing work aims to probe the mechanical response of LCEs for two limiting cases of the non -bonded interactions explored here. The first limiting case entails purely athermal chains which may undergo entropic segregation if the blocks have sufficiently disparate stiffness. Indeed, natural[64] and synthetic[125] LCE networks do not necessarily have constant stiffness across the chain length. The second limiting case entails highly attractive interactions, which can mimic ionomer chains where charged species cluster into lamellar domains with interfacial planes normal to the extensional direction[132], quite similar to the behavior seen here.

5.5 Acknowledgements

The authors thank Poornima Padmanabhan and Endian Wang for insightful discussions. This work was supported by the National Science Foundation award CMMI 1435852.

5.6 Supplemental Information

5.6.1 Simulations of systems with uncrosslinked chains

To elucidate the effects of crosslinking on morphology, we simulate the reference network at various compositions and remove the crosslinks hence resulting in a bulk system of linear 20-mer chains. We follow a similar methodology described in the main paper to prepare the systems, where the Lennard- Jones parameters are set to be all repulsive initially, and upon a relaxation period the correct parameters are enacted (consistent with $\chi N = 32$). For chains with Kbend=4.0, the morphology is lamellar for both 40%A, and 30%A, while for 20%A it shows weakly segregated, perforated lamellae. To compare to previous work, we also simulate these systems for Kbend=0. Our results match those of similar systems (in composition and χN[83]), in particular, 30%A is weakly segregated, and 20%A is disordered. This confirms that both chain stiffness and crosslinking affect the resulting microsegregated morphology.

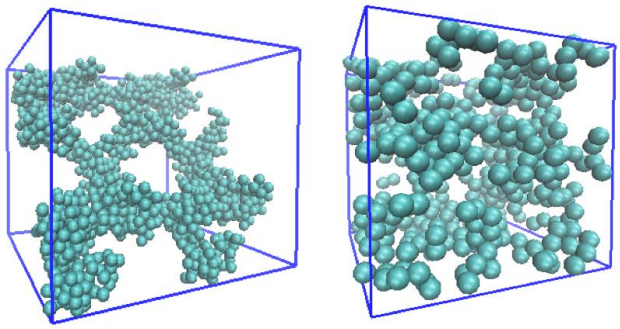

Figure 5.19: Left: morphology of 20%A, uncrosslinked ABA chains, with Kbend=4.0. Right: morphology of uncrosslinked flexible 20%A ABA chains. Only the A component beads are shown.

5.6.2 Simulations using different temperatures, and strain rates

Typical elastomers show higher Young's modulus for strain rates faster than the molecular relaxation time scale. Additionally, the mechanical response of these systems is dependent on T not only because of its effect on the inter-block enthalpic disparity (as $\chi N \propto 1/T$) and morphology, but also because of its effect on entropic elasticity. Plotted in Figure 5.20 is the tensile response of the 2×2×2 unit cell "reference network" for multiple deformation rates and temperatures. The toughness (for α from 1 to 5 as per our convention) exhibits no clear trend as strain rate is changed (Figure 5.20-left) and the saw-tooth pattern in stress is largely unaffected by strain rate (Figure 5.20-right). The lack of a systematic trend in toughness for the range of strain rates tested is a reflection of a larger variability in the values of α, and σ^* at which new domains form (and stress peaks develop). Since the formation of domains can be seen as an activated process, we conjecture that faster deformation rates allow for a faster concentration of stresses in the system (relative to the chain relaxation modes) thus facilitating the early crossing of transitional energetic barriers. Figure 5.20-left also shows an approximately linear dependence of toughness on reciprocal temperature, an expected trend for a system controlled mainly by χN [i.e., as per Eq. (5.10) in the main text]. Some peculiar non-linearities arise for $T^*>2.0$ whose origin and characteristics will be explored in a future work.

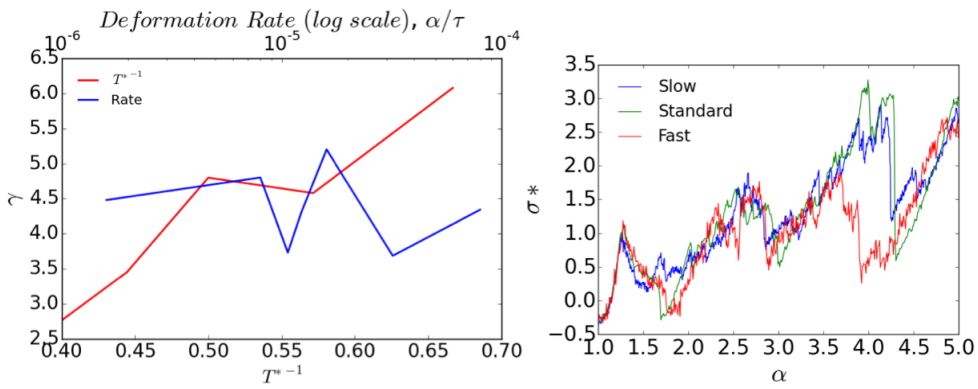

Figure 5.20: Left: Toughness of the 50%A, 2×2×2 cell reference network at various T* and deformation rates. Right: Plot of representative tensile responses for a very slow deformation rate (1.6×10^{-6}), the standard deformation rate (8.0×10^{-6}), and a fast deformation rate (3.2×10^{-5}).

5.6.3 Simulations of 2×2×2 unit cell perforated networks, and a network with symmetric, 30%A, Kbend =5.0, 40 bead chains.

Equilibration of a perforated network using a 2×2×2 unit cell lattice yielded perforated lamellae normal to the [110] direction, with the results of tensile tests plotted in Figure 5.21. Overall, the response still shows the same isotropization and the heights of the teeth are consistent with those of the larger tested network. As mentioned in the main text, a chain with twice the contour length and twice the stiffness should give a chain whose behavior is similar to that of the original 20 bead chain. Figure 5.22 shows a snapshot of the morphology as well as the tensile response of this longer chain network (2×2×2 unit cells). The x axis for the tensile response is the ratio of the simulation box length to the maximal length of the network, as opposed to α. The hairpins in the 40-mer network are twice longer than the hairpins in the 20-mer network; however, the equilibrated simulation box length is only larger by a factor of $2^{1/3}$, indicating that α is not an accurate descriptor of deformation when comparing networks of different chain lengths. The 40-mer network

shows an extra peak because the morphology has a simpler soft deformation mechanism, leading to lamellar rearrangement at smaller deformations, allowing for the existence of this previously unseen tooth. Similar values of deformation (0.15, and 0.25) are observed when a new smectic domain is formed, while the heights of peaks are different due to the increased value of χN associated with the longer chains.

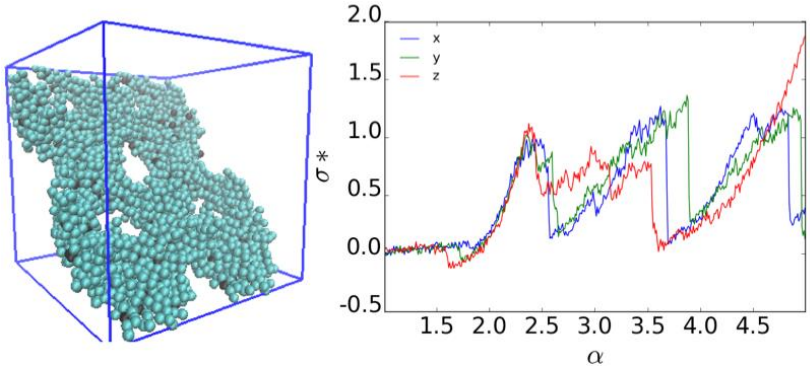

Figure 5.21: Left: Snapshot of the perforated morphology of a 2x2x2 unit cell lattice for the perforated network. Right: Plot of a representative tensile response. Blue beads are A type, black beads are cross-links, and the B block was removed for clarity.

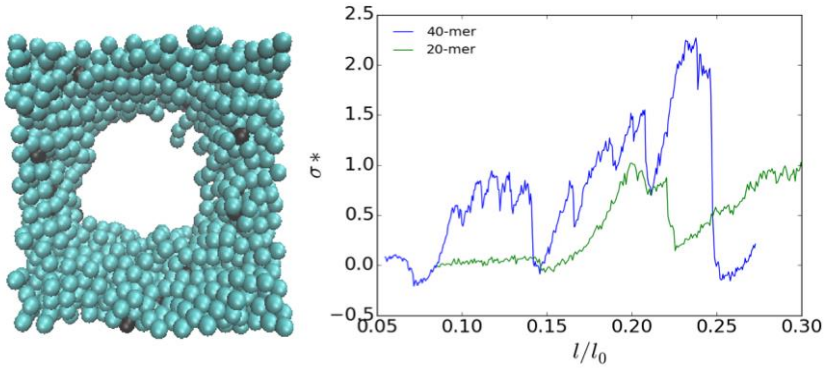

Figure 5.22: Left: Snapshot of the untilted, perforated, lamellar morphology obtained for the 40-mer network. Blue beads are A type, black beads are cross-links, and the B block was removed for clarity. Right: Plot of the tensile response for this network and the 2×2×2 network, normalized to the maximal extension of the respective networks.

6 Effect of Block Immiscibility on Strain-Induced Microphase Segregation and Crystallization of Model Block Copolymer Elastomers

6.1 INTRODUCTION

Main-chain liquid crystalline elastomers (LCE) are interconnected networks of polymer chains that are semi-flexible, i.e., they undergo nematic ordering under certain stimuli[34,44,124,188,204]. Many applications for LCE have been considered such as high-strain actuators[218,232] and artificial muscles[125,222]. These applications leverage the coupling between the macroscopic shape of the material and the microscopic orientational ordering. When the chains align, their prolate shape leads to anisotropic change in macroscopic shape of the material. Along with this coupling, nonlinear elastic behaviors, such as near-zero stress deformations, are observed at both macroscopic and microscopic levels for natural LCEs[95,225,228] and synthetic LCEs[211]. Macroscopically, the "goldstone modes" are observed[153], where the direction of predominate chain alignment (henceforth called "the director") can reorient with zero-stress deformations. When a deformation is applied to LCE along a direction misaligned with the director, the chains can reorient while the cross-links can rearrange to leave unchanged their overall distribution (and entropy). Microscopically, near-zero stress deformations are seen due to the characteristic way semi-flexible chains deform. For instance, a common chain configuration is that of a hairpin, and the mechanism of chain unfolding entails "shifting" the mass of one hairpin leg to the other leg until the first leg is short enough that the bending energy overcomes the chain's tendency to stay aligned with its neighbors and fully unfolds[226]

(Figure 6.1a). This shifting of mass from one hairpin leg to the other occurs largely at zero change in entropy due to minimal transverse chain motion, resulting in zero stress.

LCEs made of semi-flexible chains with their ends connected by tetra functional cross-links, and with no entanglements (Tetra-LCE, T-LCE) have been shown to exhibit a "saw-tooth" stress response to uniaxial deformation. This behavior leverages both the tendency to form hairpins and the coupling between chain alignment and material shape. Due to main-chain backbone semi-flexibility, the chains and cross-links self-assemble into separate layers (Figure 6.1)[4,18]. Upon deformation, stress builds in these networks until chain backbone hairpins unfold and a new cross-link layer forms. Upon hairpin unfolding, the network length increases, reducing chain overstretching. When the constituent chain in the T-LCE are ABA block copolymers (henceforth referred to as BCP), an additional interfacial energy penalty enhances the segregation, and leads to the formation of smectic layers (SL), namely, orientationally and translationally ordered domains of A and B blocks[4,149] (Figure 6.1b). While forming a new domain during network deformation, both interfacial area and energy increase, requiring larger stresses for SL formation as compared to homopolymer T-LCE. A succession of large stress peaks occurs upon further deformation (Figure 6.1c) reminiscent of naturally occurring tough materials such as titin and spider silk, both of which undergo non-affine domain deformations to relieve stress[18,91,109,148,171,230]. This saw-tooth behavior increases the toughness, γ, the energy absorbed during deformation, of the T-LCE. As such, these materials have potential applications as reinforcing fibers, adhesives, or shock absorbers by allowing greater energy absorption through yielding, which is the cornerstone of the "crumple zone", a concept commonly used in cars and protective equipment[79,100].

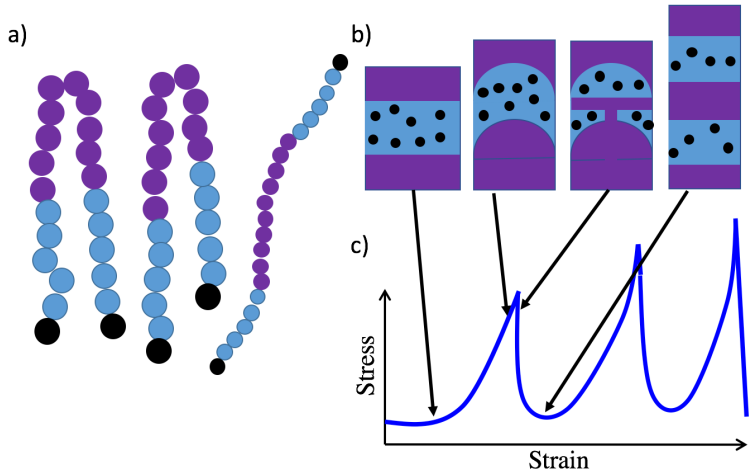

Figure 6.1: a) Depiction of unfolding mechanism for semi-flexible chains (colored as block copolymer for clarity). Mass on one side of the hairpin shifts until the bending energy causes the unfolding. b) Cartoon of the saw-tooth mechanism in T-LCE with block copolymer chains, pointing to a representative position on the stress-strain cure (c). *From left to right:* Initial system stays at zero stress upon straining by removing any "slack"; morphology begins to bend incurring in an interfacial energy penalty and stress buildup; parent layer begins to break up after a greater amount of strain as the hairpins unfold; two daughter layers form from the parent layer, relaxing the stress.

Previous work has established the synthetic viability of these materials by showing that the saw-tooth behavior is present even in networks that were created through a simulated synthesis[152]. Since the tensile behavior of such realistic LCEs is largely captured by the T-LCEs (of identical chain length), T-LCE can be used as a convenient model to test the effect of different chain chemistries/architectures, as was previously done to find new deformation behaviors[149]. Since γ for T-LCE is nearly an order of magnitude larger for ABA BCP chains (for blocks in the strong segregation regime) than for homopolymer chains, enthalpic effects associated with inter-block interactions are the dominant factor. We can then examine the effects of altering χN, a parameter quantifying both block immiscibility and reciprocal temperature, on the macroscopic and microscopic response to uniaxial strain in T-LCEs. If the interfacial energy is the dominant factor in our model T-LCEs, γ would be expected to be

linearly dependent on χN, which holds if the deformation behavior (i.e., the number and position of peaks in the saw-tooth response) is unchanged across different χN and the system is above its melting and glass transition temperature. However, at high χN crystallization may ensue, a process that would involve an additional resistance to deformation. The concept of using polymer crystals for mechanical strengthening is not new and occurs, e.g., in spider drag-line silk, where the crystalline regions act as strengthening crosslinks, and as the regions where domain yielding occurs to give super-toughness quality to the silk.

Our BCP T-LCEs embody 3 key design elements: semi-flexible chains, a network topology, and a BCP chain architecture. These elements arguably work in concert to provide a high-toughness saw-tooth deformation behavior depicted in Figure 6.1 that is distinct to that of other classes of polymeric materials. Flexible homopolymer chains, semi-flexible homopolymer chains, and networks of flexible homopolymer chains all lack the ability to form the segregated domains associated with BCPs. Networks of fully flexible BCP were found to lack the "teeth" in the tensile response[152] due to the absence of the microscopic orientational ordering characteristic of T-LCE. Indeed, chain stiffness promotes orientational alignment that facilitates chain mobility along the director, the unfolding of hairpins, and thus localized network deformations. Homopolymer T-LCE were shown to exhibit features of the saw-tooth behavior[18], but due to the lack of interfacial energetic penalties the formation of layers of cross-links is not concerted and leads to much weaker saw-tooth peaks and hence a lower toughness. Given the dominant contribution of χN to γ for T-LCEs, how important is crosslinking on the stress response to uniaxial deformation? i.e., what would the response be if BCP T-LCE with a lamellar conformation were "uncrosslinked" before deformation? We use a theoretical analysis and molecular simulations to elucidate how the

crosslinking, which couples the lamellar layers, promotes the distinct saw-tooth stress response.

The rest of the paper is organized as follows. In Section 6.2 we describe our simulation model and methods. The results in Section 6.3 is broken into two parts where in section 6.3.1 we explore the impact of χN on the deformation of non-network BCP to gain a better understanding of the factors influencing the saw-tooth tensile response seen in BCP T-LCE. In Section 6.3.2 we examine the influence of χN on the deformation of BCP T-LCE and how the emergence of crystalline regions affects the deformation behavior. In Section 6.4 we summarize our findings and propose new avenues to explore.

6.2 MODEL AND METHODS

The network consists of tetra-functional cross-links that sit on lattice points of a diamond unit cell when the network is fully expanded (see Figure 6.2). Each network consists of 512 cross-links and 1024 chains connected across the periodic boundaries to simulate an infinite network. The diamond connectivity results in an entanglement free network, key to the mobility of the cross-links and the creation of the observed smectic domains[4]. The chains are monodisperse 20-mers (unless otherwise stated), of coarse-grained Lennard-Jones (LJ) type beads chained into a symmetric ABA architecture [f_A/2-(1- f_A)- f_A/2] where f_A is the volume fraction of the A-type block. To work in reduced LJ units, reported properties are scaled as follows: $T^*=Tk_b/\varepsilon$, $\tau=t(\varepsilon/(m\sigma_{LJ}^2))^{1/2}$, $\sigma^*=\sigma\sigma_{LJ}^3/\varepsilon$, $P^*=P\sigma_{LJ}^3/\varepsilon$, $\gamma=\Gamma\sigma_{LJ}^3/\varepsilon$, and ρ=(# of particles)/ σ_{LJ}^3, where T is temperature, σ^* is stress, P is pressure, Γ is the toughness of the material, ρ is number density, and all scaling variables (σ_{LJ}, m, ε) are set to 1.

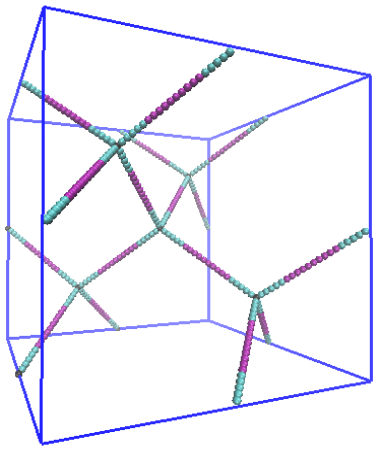

Figure 6.2: Fully swollen unit cell with 16, 20-mer chains and 8 cross-links in an ABA arrangement. The A, B, and cross-link beads are colored blue, purple, and black, respectively.

The potentials used in this work are chosen to represent a generic coarse-grained polymer behavior, closely following the Kremer-Grest model[71], beginning with bonding that is modeled by the standard FENE potential:

$$U_{FENE} = -KR_0^2/2\ln\left(1-\left(\frac{r}{R_0}\right)^2\right)+4\varepsilon\left(\left(\frac{\sigma_{LJ}}{r}\right)^{12}-\left(\frac{\sigma_{LJ}}{r}\right)^{6}+\frac{1}{4}\right) \quad (6.1)$$

where the LJ-like potential term is cut off at $r=r_c=2^{1/6}\sigma_{LJ}$ to mimic excluded volume interactions and $K= 30\varepsilon/\sigma_{LJ}^2$ and $R_0 = 1.5\sigma_{LJ}$ are the spring force constant and maximum bond extension length, respectively. These widely used values of K and R_0 describe a stiff bond whose equilibrium length is approximately equal to the bead diameter ($\sim 0.97\sigma_{LJ}$); such choices are not very influential on the properties of interest in our study. For non-bonded interactions, a cut and shifted LJ potential is used with the form:

$$U_{LJ} = 4\varepsilon\varepsilon^{ij}\left(\left(\frac{\sigma_{LJ}}{r}\right)^{12}-\left(\frac{\sigma_{LJ}}{r}\right)^{6}-\left(\frac{\sigma_{LJ}}{r_c}\right)^{12}+\left(\frac{\sigma_{LJ}}{r_c}\right)^{6}\right) \quad (6.2)$$

for $r/\sigma_{LJ} \leq r_c$ ($U_{LJ} = 0$ otherwise), where $r_c^{AA} = r_c^{BB} = 2.5\sigma$ (attractive potential), and $r_c^{AB} = 2^{1/6}\sigma$ (repulsive potential); A blocks and cross-links are taken to be the same type. Table 6.1 lists

the dimensionless ε^{ij} values (reduced with respect to ε) employed for all interaction pairs and the corresponding effective Flory-Huggins parameter χN.[83] These χN values are based on relationships derived for flexible di-block chains rather than the tri-block semi-flexible chain network used here; however, they are expected to capture well the relative trend.

Table 6.1: Approximate mapping between the simulation model parameter ε^{ij} and corresponding Flory's interaction parameter and effective temperature (relative to ODT temperature).

$\boldsymbol{\varepsilon^{ij}}$ (ε units)	$\boldsymbol{\chi N}$	$\boldsymbol{T^{*}_{eff} = \varepsilon^{ij}_{ODT}/\varepsilon^{ij}}$
0.25	15	0.725
0.50	32	0.362
0.75	47	0.242
1.0	63	0.181

Our model deviates from the Kremer-Grest model in that mesogenic behavior is induced by a bending potential that is applied to all non-cross-link beads, which has the form:

$$U_{angle} = k_b T K_{bend}\left(1+\cos(\psi)\right) \qquad (6.3)$$

where K_{bend} is a stiffness constant and ψ is the angle between two consecutive bond vectors. For all our simulations we use a constant value of T^{*}= 2.0 which will set the chain length of a given polymer needed to match a prescribed χN of the model. It is important to note, however, that since the way we prescribe χN is by simultaneously rescaling ε^{ij} for all non-bonded LJ interactions, then varying χN is essentially equivalent to changing the effective temperature of the system T^{*}_{eff}, whose value can be defined relative to the order-disorder

transition temperature (ODT) of our model network (i.e., $T^*_{\mathrm{eff}} = 1$ at the ODT), and is approximately given by:

$$T^*_{\mathrm{eff}} = \frac{\chi N_{\mathrm{ODT}}}{\chi N} \approx \frac{\varepsilon^{\mathrm{ij}}_{\mathrm{ODT}}}{\varepsilon^{\mathrm{ij}}} \quad (6.4)$$

From previously derived relationships[83] we estimate χN_{ODT}=10.5 and $\varepsilon^{\mathrm{ij}}_{\mathrm{ODT}} = 0.18$, and the values of T^*_{eff} corresponding to the $\varepsilon^{\mathrm{ij}}$ values used are also given in Table 6.1. Note that T^*_{eff}, unlike the true temperature, does not affect U_{angle}. While either $\varepsilon^{\mathrm{ij}}$, χN, or T^*_{eff} could be used to define block immiscibility, henceforth we will only use $\varepsilon^{\mathrm{ij}}$ as it is the parameter being directly changed in our simulations; however, the mapping given in Table 6.1 should help translating $\varepsilon^{\mathrm{ij}}$ values into physically meaningful quantities.

We purposely chose a generic polymer model whose potential functions and parameter values are based on well-established models because it is well-suited to capture physical trends of polymer behavior, but also because it allows us to build on the existing body of work by studying systems with added features of molecular complexity. Indeed, we build on previous studies that have used similar models to describe flexible homopolymers, BCP[83], athermal rigid chains[55], and networks.

To prepare the networks for tensile tests, a multistep procedure is implemented to concurrently generate configurations for all 4 $\varepsilon^{\mathrm{ij}}$ values of interest. The preparation begins with a fully swollen network with $\varepsilon^{\mathrm{ij}}$=0.25. That system is compressed in two stages: a compression to a density of ρ=0.5 in 10^7 time steps, followed by a slow compression to ρ_{melt} in 5×10^7 time steps. The compression follows the rate law $l = l_0 \exp(k\tau)$, where l is the simulation box length, l_0 is the original box length, k is a dimensionless compression rate, and τ is the scaled LJ time. The slow compression from ρ=0.5 to ρ_{melt} allows for faster

equilibration as chains can rearrange more easily. Once the network is compressed to ρ_{melt} and equilibrated, it is subjected to the following cycle. In the up-cycle, the value of ε^{ij} is gradually increased to the next value (0.25→0.5→0.75→1.0) over 2×10^7 timesteps and allowed to equilibrate for another 2×10^7 timesteps. We consider a system equilibrated when steady state values are reached for the bending, non-bonded, and bonded potentials, as well as the pressure tensor components, p_{xx}, p_{yy}, and p_{zz}. This is repeated until ε^{ij}=1.0 and equilibrated, where the cycle is reversed (down-cycle) where ε^{ij} is decreased to the next value (1.0→0.75→0.5→0.25) following the same scheme as the up-cycle (during both cycles, ρ is kept constant). The morphology, values for bending, bonded, and non-bonded energy as well as the pressure tensor components for each value of ε^{ij} are the same during up-cycle and down-cycle. This suggests that our morphologies (all lamellar) are not greatly affected by our choice of equilibration scheme. A snapshot is taken for each value of ε^{ij} after the appropriate equilibration in the down-cycle, and a modified NPT ensemble equilibration run is performed on each of these snapshots. In this equilibration, the directions perpendicular to the lamellar plane normal are coupled (i.e., their dimensions vary in unison) while the pressure component along the plane normal varies independently. The pressure is set as the ensemble average pressure from the NVT equilibration in the directions perpendicular to the lamellar plane normal (i.e., if the lamellar normal points in the x direction, then the set pressure is $P^*_{set} = 0.5\left(\langle P^*_{yy}\rangle + \langle P^*_{zz}\rangle\right)$. During this equilibration, the box expands along the plane normal and contracts in the perpendicular directions. P^*_{set} is adjusted so that the equilibrated density is within 0.84-0.86. The value of P^*_{set} which achieves the target density range for a given ε^{ij} is different (see Table 6.2 for specific values). The final box length along the lamellar plane normal (the axis which is deformed) increases as ε^{ij} increases. This scheme

is used because, as reported in previous studies, the kinetics of these networks are slow, and combined with the high energetic barriers associated with the larger values of ε^{ij} would lead to kinetically arrested states. Attempts to equilibrate systems where ε^{ij}=1.0 without the annealing cycle described resulted in kinetically trapped, disordered cocontinuous morphologies. The cycling methodology allows for networks with ε^{ij}=1.0 to have a lamellar morphology. Additionally, by linking the morphologies across different values of ε^{ij} (i.e., the ε^{ij}=1.0 morphology is based on the ε^{ij}= 0.25 morphology) the variability due to large discrepancies in starting configurations is removed, and the effect of changing ε^{ij} is isolated.

In all simulations, the temperature and pressure are controlled using a Nosé-Hoover thermostat/barostat (depending on ensemble), with T^*=2.0 and P_{ii}^* being set according to the procedure described above. The equations of motion are integrated using a velocity Verlet algorithm, with a timestep equal to 0.005τ.

The fully equilibrated system is considered the unstrained α=1 state, where α is the extension ratio, i.e., the length of the box edge along the deformation direction scaled by the undeformed plane normal box length (l_{ext}/l_0). The deformation direction is always taken along the lamella plane normal to minimize the occurrence of soft deformations associated with goldstone modes. To simulate uniaxial deformations, the box length along the deformation axis is expanded linearly in a strain-controlled ensemble every 5000 time steps up to α=5 in 2×10^8 time steps, leading to a deformation rate of $4\times10^{-6}\,\alpha/\tau$, while the other two dimensions are coupled and controlled using a barostat set to P_{set}^*. The stress is calculated using:

$$\sigma^* = -P_{ii}^* + \lambda\left(P_{jj}^* + P_{kk}^*\right) \qquad (6.5)$$

where P_{ii}, P_{jj} and P_{kk}, and λ are the extensional direction pressure tensor component, the non-

extensional direction pressure tensor components, and the Poisson ratio, respectively. When plotting σ^*, to reduce noise, we average the instantaneous value of λ across the entire deformation and use this as a constant value for λ in Eq. (6.5). To quantify the chain alignment in the system, we calculate the bond order parameter P_2, which is defined as the largest eigenvalue of the matrix:

$$Q_{ij} = \frac{1}{N}\sum_{n=1}^{N}\frac{3}{2}\hat{u}_{in}\hat{u}_{jn} - \frac{\delta_{ij}}{2} \quad (6.6)$$

where $\hat{u}_{in}$ is unit vector pointing along bond n in the i direction, N is the number of bonds in the system, and δ_{ij} is the kronecker delta. γ is calculated using:

$$\gamma = \int_1^{\alpha_f} \sigma^* H\left(\sigma^*\right) d\alpha \quad (6.7)$$

where $H(\sigma^*)$ is the Heaviside function, which is introduced to ignore negative stresses and thus isolate the total energy absorbed (i.e. total work done on the system); the energy "returned" (work done by the system) when stress is negative would essentially be dissipated. Conventionally, the upper limit of integration in Eq. (5.7) should be α at fracture; however, since our model bonds do not account for breaking events, α= 5 was chosen for convenience. To provide some calibration for the toughness values to be reported, we can compare the simulated toughness of ~0.15GPa[217] (calculated for extension from α=1 to 5) for an atomistic model of a polyethylene network with diamond connectivity to γ=0.33 for a fully flexible coarse-grained homopolymer network of comparable molecular weight; hence a unit in γ in our model would correspond to ≈0.4GPa. Repeated runs of the same network with differing starting configurations give slightly different results. As such, we show representative plots and report the average values where possible. Stress-strain plots show values averaged over ~0.005-0.01α.

6.3 RESULTS AND DISCUSSION

6.3.1 Role of Block Immiscibility on Uniaxial Deformation of Non-Network BCPs

A first step toward elucidating the role of ε^{ij} on tensile behavior of BCP T-LCEs is to understand the uniaxial deformation behavior of uncrosslinked BCP chains. The case of fully flexible chains can be studied by considering the theoretical changes in free energy of a lamellae-forming BCP confined between two plates as the inter-plate separation is varied at constant volume according to the following equation (Eq. (4) in Ref. [216] for a system interacting with the plates as if they were bulk material):

$$\frac{F}{F_0} = \frac{1}{3}\left(\left(\frac{\alpha}{n_l}\right)^2 + \frac{2n_l}{\alpha}\right) \quad (6.8)$$

where n_l is the number of A-block layers between the plates. Figure 6.9 in the SI shows a plot of F/F_0 over a range of n_l, and deformations relevant to our tensile experiments. The envelope of the minima in the manifold of free energy curves shows that n_l increases as α increases where the average deformation needed to increase n_l remains fairly constant, in qualitative agreement with the tensile behavior seen in T-LCE, and with the results in Ref. [216]. This suggests that uniaxial deformation of a lamellar BCP could produce a sequence of small stress teeth. However, our uniaxial deformation simulations of both flexible and semi-flexible BCP melts reveal that lamellae tilting, rather than new domain formation, is the primary mechanism for maintaining the domain spacing constant (see Figure 6.10), consistent with experimental observations[33,41]. In larger systems, lamella tilting need not occur uniformly across the whole sample but heterogeneously across numerous grains. The free-energy calculations based on Eq. (6.8) ignores lamella tilting as the more effective mechanism to minimize any free-energy penalty in deformed BCP melts (provided that lamellar orientation is mechanically decoupled from the strain axis).

The above analysis singles out network connectivity as a needed factor in producing the saw-tooth behavior observed in BCP T-LCE as depicted in Figure 6.1[4,149] (with the other 2 factors being backbone stiffness and BCP segregation as alluded to in the Introduction). The BCP nature provides an energetic driving force that favors the concerted formation of lamellar domains, but it is the network connectivity that provides the coupling between domains and the strain field to favor the lamellae orientation perpendicular to the deformation direction.

6.3.2 Role of χN on the Toughness of BCP T-LCEs

This Section examines the large and fundamental role played by χN, as described by ε^{ij} in our model, on the magnitude of γ. Figure 6.3 shows how γ depends on A-block volume fraction (f_A) and ε^{ij}. We observe that for any f_A, γ always increases with ε^{ij}, but for any given ε^{ij}, γ follows a non-monotonic trend with f_A. Specifically, for $\varepsilon^{ij} \geq 0.75$, γ depends on f_A following a roughly parabolic profile with a maximum at f_A=0.5, a result consistent with those previously reported, while for $\varepsilon^{ij} \leq 0.5$, γ exhibits little or no compositional dependence on f_A.

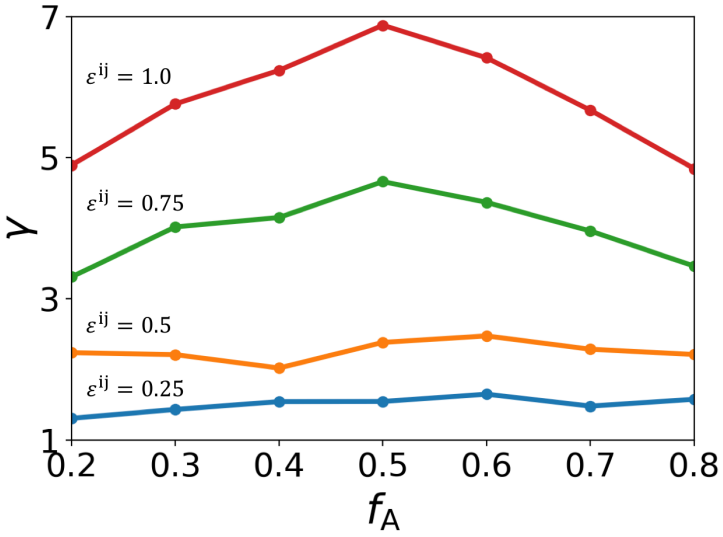

Figure 6.3: Toughness (γ) as a function of block composition (f_A) for various values of ε^{ij}.

To understand the irregular compositional dependence for $\varepsilon^{ij} \leq 0.5$ in Figure 6.3, we calculated the change in the interfacial surface area throughout the deformation. Previously, the non-bonded interactions between A and B (i.e., surface tension) were assumed to be the primary contributor to the stress peak associated with the process of domain formation. As such, a difference in the interfacial area as a function of f_A would lead to the main compositional dependence of γ. No significant difference in surface areas is observed between compositions when $\varepsilon^{ij} \leq 0.5$ (Figure 6.11 in the SI), consistent with Figure 6.3. In contrast, for $\varepsilon^{ij} \geq 0.75$ a difference in the stress peaks is observed after $\alpha \gtrsim 2.5$, as shown in Figure 6.4a for two representative tensile curves for $f_A = 0.2$ and 0.5, and $\varepsilon^{ij}=1$. Analysis of the local values of the P_2 (nematic order parameter) for the network bonds shows that the onset of f_A dependence for systems with $\varepsilon^{ij} \geq 0.75$ correlates with the observation of high P_2 values (≥ 0.85). These large P_2 values are due to partial crystallization of the chains into a hexagonal lattice (Figure 6.4c). We note that crystallization of LCEs has been observed experimentally

in the regime of both temperatures and strains we use here,[106] consistent with our observations.

Crystallization in our systems is not only induced by large ε^{ij} (corresponding to low temperatures) but also by strain: below a certain threshold crystals do not form even for $\varepsilon^{ij} \geq 0.75$. Figure 6.5 shows the crystal fraction for each value of ε^{ij} as a function of α, averaged over all compositions, and a sequence of snapshots illustrating the development of crystalline domains. The crystal fraction is defined as the number of atoms which are part of at least one bond with a local value of $P_2 \geq 0.8$, divided by the total number of atoms.[133,233] As a complement, we also assigned crystal character to atoms with high local bond order parameter q_6 as described in Ref. [206]. Unlike P_2 that measures local orientational order of bonds, q_6 measures local translational order with the hexagonal-lattice packing (as per Figure 6.4c). Figure 6.5(b) shows that the two metrics of crystal fraction are consistent and indicate that a distinct increase in order occurs around α=2.5-3, coincident with the conditions where the compositional dependence was found to emerge in Figure 6.4. Further, both metrics display a saw-tooth pattern where crystallinity tends to increase with deformation but to decrease upon domain formation. This is consistent with the microscopic deformation mechanism previously elucidated (Figure 6.1). In fact, as the network is strained macroscopically, the chains tend to stretch, increasing both the stress and chain alignment; however, upon the formation of a new lamellar domain, the chains can relax and the crystal fraction and stress drop. Some discrepancy between the P_2 and q_6-based crystal fractions is seen in the curves for $\varepsilon^{ij} = 0.25$ at any α and for all ε^{ij} at small α, where the q_6–based crystal fraction is larger, likely due to a lower sensitivity of q_6 to discriminate against disorder.

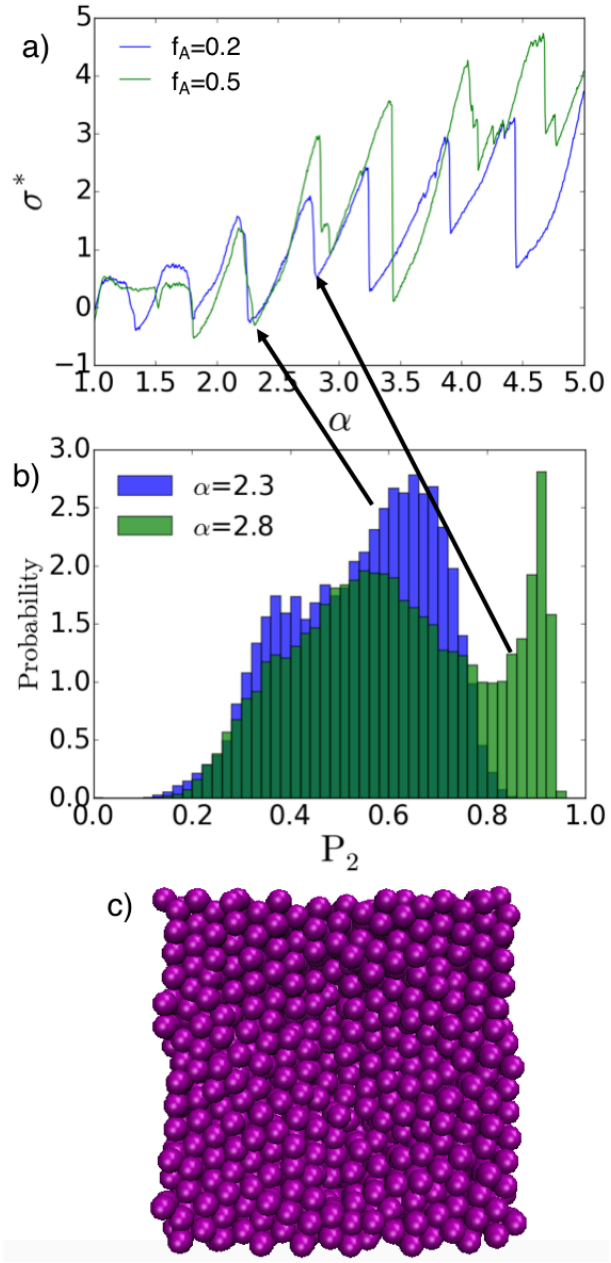

Figure 6.4: a) Tensile plot for two T-LCE (f_A = 0.2 and 0.5) with ε^{ij}=1.0. b) Plot of local P_2 values for the f_A = 0.2 network at α=2.3/2.8 with arrows point to where they correspond on a). c) Simulation snapshot of the network cross-section for a partially crystallized network showing hexagonal packing.

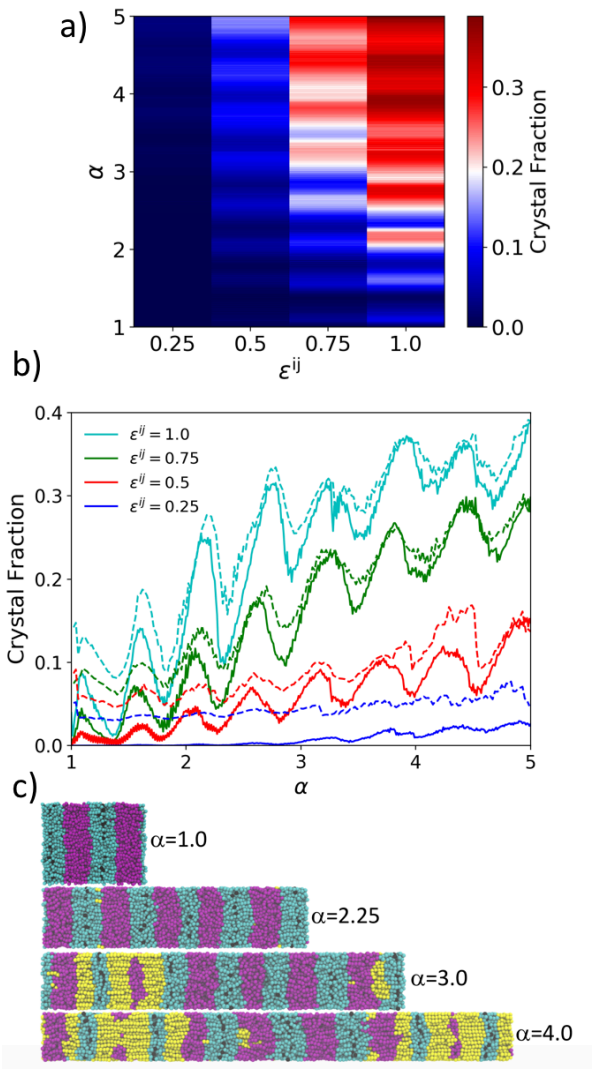

Figure 6.5: (a) Heat map of crystal fraction as a function of a and ε^{ij}. (b) Plot of crystal fraction as a function of α for each value of ε^{ij}. (a) and (b) show results averaged over networks of all compositions. Solid lines and dashed lines represent the results based on P_2 and q_6 order parameters, respectively. (c) Snapshots of a system (f_A = 0.5, ε^{ij}=1.0) at various deformations, coloring in yellow crystalline beads (determined using P_2 metric).

Since crystallization can be driven by both high ε^{ij} and high strain, is crystallization for $\varepsilon^{ij} \leq$ 0.5 systems completely absent or is it still present but at much larger deformations (i.e., α> 5)? To address this question, a DBC T-LCE (f_A = 0.5, ε^{ij}= 0.25) was deformeqd to maximal extension (before stress divergence). The observed tensile response (Figure 6.6a) and

crystal fraction (Figure 6.6b) showed no evidence of crystallization for α<9, where the crystal fractions after domain formations ($P_2 \approx 0.05$) are well below those for $\varepsilon^{ij} \geq 0.75$ system ($P_2 \approx$ 0.2-0.3 for whole system). Large crystal fractions in the $\varepsilon^{ij} = 0.25$ system are only observed as the network approaches its limiting extension ($a \rightarrow 10$) when all chains become hyperstretched and aligned. This minimal tendency to crystallization for $\varepsilon^{ij} \leq 0.5$ systems is expected because decreasing ε^{ij} implies larger "effective" temperature in our systems as stated by Eq. (6.4), which makes the amorphous (melt) state more resilient to strain-induced ordering.

Having established that a high ε^{ij} is crucial for the emergence of crystallization, we address now the related question: how important is the magnitude of the interfacial energy (between A and B blocks)? Namely, would a "homopolymer" network (i.e., having only one block where all interactions are attractive and $\chi N = 0$) show a similar tendency to crystallize under strain? To isolate the effect of interfacial energy, deformation simulations to maximal extension were performed for two systems: a DBC T-LCE ($f_A = 0.5$, $\varepsilon^{ij}=1.0$) and a "homopolymer" system. The latter is identical to the DBC T-LCE system in all respects (i.e., same ε^{ij}, temperature, initial starting configuration, P^*_{set}, and composition) except that $r_c^{ij}=2.5\sigma_{LJ}$ for all ij pairs, which makes all interactions effectively like-pair interactions, hence removing any effect of microphase segregation and the presence of block interfaces. The tensile responses (Figure 6.6a) and crystal fractions (Figure 6.6b) for these systems suggest that the observed crystallization is influenced by both the strength of like-pair interactions and the interfacial energy. In fact, Figure 6.6(a) and Figure 6.6(b) indicate that while interfacial energy plays a lesser role than a strong like-pair interaction strength (low temperature) in bringing about crystallization, it plays a dominant role in the enhancement

of the saw-tooth pattern and associated toughness. As found in former studies,[4,149] here we also observe that both systems exhibit multiple transitions associated with the formation of domains where chains and crosslinks segregate giving rise to a sawtooth stress pattern. Compared to the DBC T-LCE, the homopolymer network exhibits a delayed onset of crystallization (at $\alpha \approx 4$ compared to $\alpha \approx 2.5$), and a partial melting of the crystalline region after $\alpha \approx 7$ (snapshots of the homopolymer system with the crystalline domains colored are shown in Figure 6.6d). These trends in the homopolymer network are explained by the lack of interfacial energy which allows for easier movement of cross-links between chain layers, so that crystal domains are more easily disrupted by the formation of new domains. This difference in the ability for cross-links to move between the smectic layers is shown in Figure 6.6c where the cross-link density distribution along the deformation direction is plotted for both the homopolymer and DBC networks with $\varepsilon^{ij}=1.0$. The density distribution for the homopolymer network shows lower variability in the density of cross-links than the DBC system, indicating that the cross-links in the DBC system are not as free to move between layers as the homopolymer network.

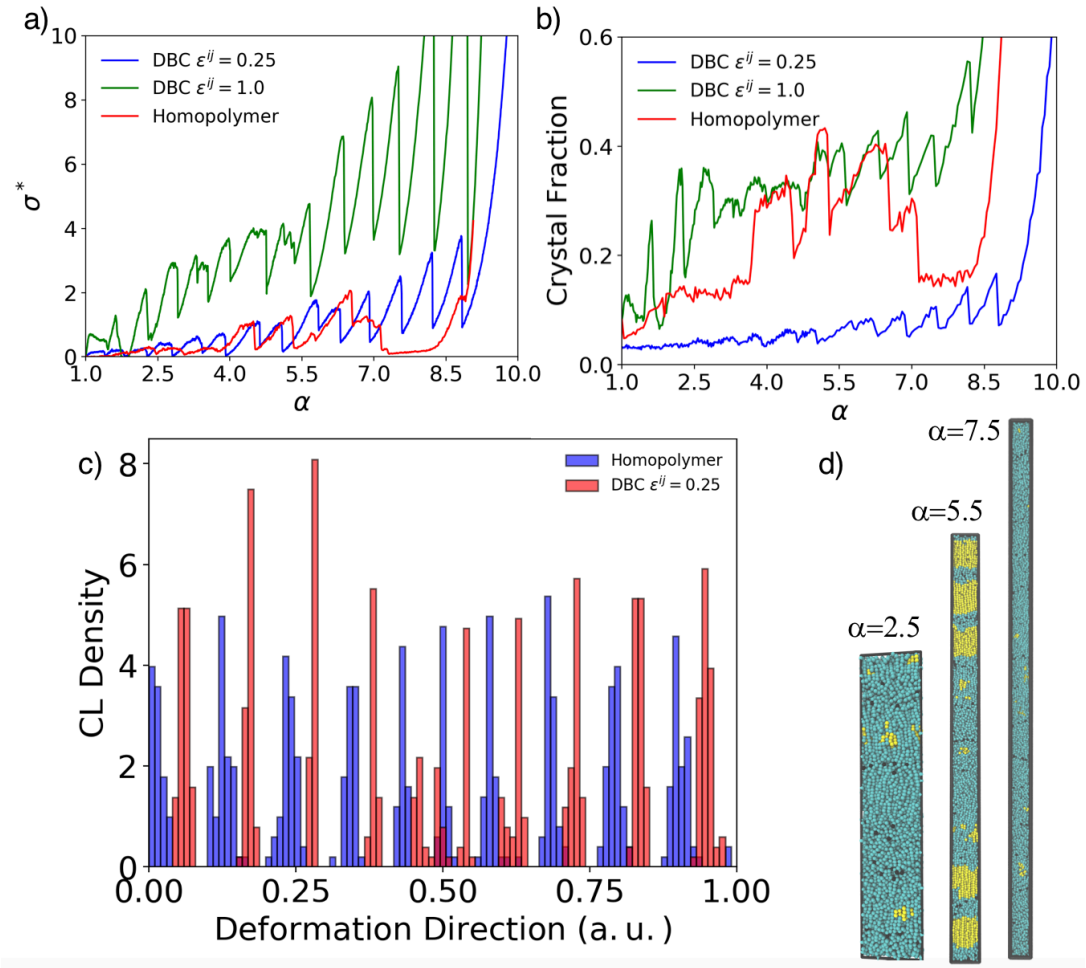

Figure 6.6: a) Tensile responses for f_A = 0.5 DBC T-LCE networks with ε^{ij}= 0.25, and 1.0, and a homopolymer network, identical to the DBC T-LCE network with ε^{ij}= 1.0 but with $r_c^{ij}=2.5\sigma_{LJ}$ for all ij pairs. b) The corresponding crystal fractions as a function of strain. c) Density distribution of the cross-links (CL) for the homopolymer and DBC, ε^{ij}= 1.0 systems for α=5.5. d) Selected snapshots of the homopolymer network with crystalline regions colored yellow (based on P_2 order parameter).

Having gained a more detailed understanding of the multiple factors affecting crystallization in our systems, we now propose a microscopic model, depicted in Figure 6.7a, to explain the emergence of the compositional dependence upon chain crystallization. This model also accounts for the lack of compositional dependence when $\varepsilon^{ij} \leq 0.5$, a behavior that contrasts previous work[1]. When a network is not partially crystallized the SL can deform into morphologies such as those seen in Figure 6.7b (top). In contrast, when the chains are

crystallized, the chains needed to form a new SL protrude from the "donor" SL as shown in Figure 6.7b (bottom). This difference in the deformation behavior can be detected by tracking the bead positions from simulation snapshots just prior to the formation of a new SL, to produce a 1-D probability density profile, $d(x)$, of the minority phase along the deformation axis as shown in Figure 6.7c (with bins along x axis of size σ_{LJ}). A key quantity of interest is the minimum probability of observing the minority phase between the two donor layers, $MPDL = \min_{x_A \le x \le x_B} d(x)$, where $x_B > x_A$ are the coordinates of the centers of mass of the two donor smectic layers, after correcting for periodic boundaries. Averaging over many "teeth" and replicate tensile tests gives approximate values of MPDL of 0.003 for $\varepsilon^{ij} \le 0.5$ and 0.001 for $\varepsilon^{ij} \ge 0.75$. These results are consistent with Figure 6.7a: a crystalline network should have a lower value of MPDL as there is no (or only sporadic) bridging between the two layers, while in a non-crystalline network the bridging should lead to a greater MPDL. Thus, prior to new SL formation the non-crystalline domains are deforming more extensively compared to the crystalline domains. A full list of values of MPDL is given in Table 6.3.

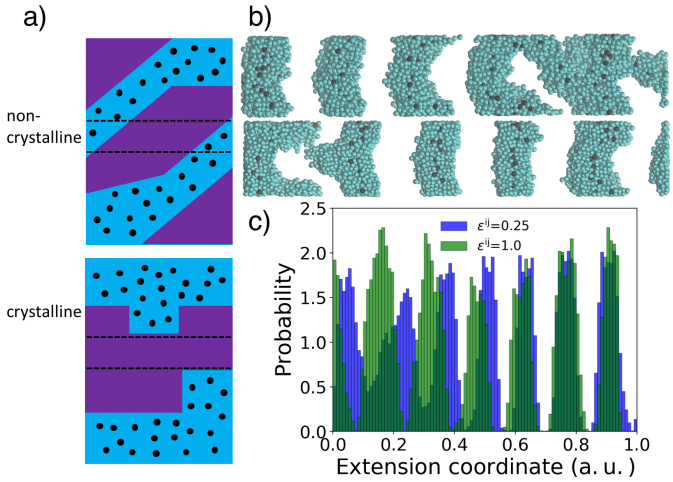

Figure 6.7: a) Cartoon representation of the two morphologies seen when the deformed network is non-crystalline (top), and semi-crystalline (bottom). b) simulation snapshots of T-LCE with f_A= 0.30, at a=3.56 for ε^{ij}= 0.25 (top), and ε^{ij}=1.0 (bottom). c) Minority block density along the extensional direction for ε^{ij}= 0.25 and 1.0.

For deformations of crystalline domains, the newly created interfacial area is expected to be proportional to the volume fraction of the minority block, as illustrated in the idealized model shown in Figure 6.13. Accordingly, the toughness can be modeled by $\gamma = C_1 \times \min(1-f_A, f_A) + C_2$, where the C's are constants, the first term is the contribution from the interfacial energy, and the second term accounts for the additional "friction" that a chain must overcome when moving in a crystalline region, and also from the contribution at small deformations where there is no crystalline material. Figure 6.8 shows γ normalized by the γ value at f_A = 0.5, where the proposed composition dependence is the line labeled "modified theory" to distinguish it from our "original theory" proposed[1] where $\gamma = C\, f_A(1-f_A)$. The latter is a model based on the assumption that γ is proportional to energetic penalty associated with the average number of A-B contacts that would occur as blocks mix while producing new lamellar domains. It is apparent that the modified theory provides a better fit to the simulation data for $\varepsilon^{ij} \geq 0.75$, giving further support to our analysis.

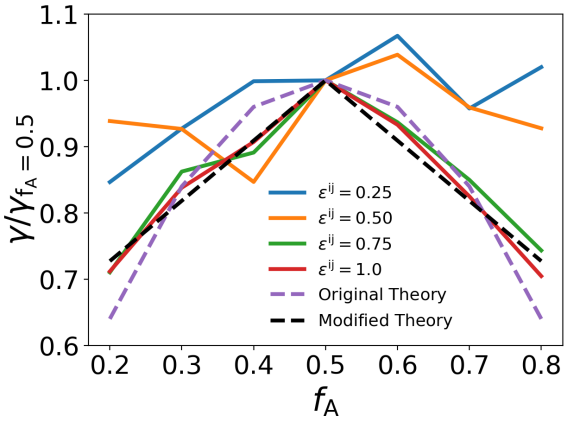

Figure 6.8: Plots of γ normalized by $\gamma_{fA=0.5}$) against f_A for various values of ε^{ij}. The predictions from our original theory[1], and the modified theory from this work are also plotted.

The lack of compositional dependence for non-crystalline systems can be correlated with the type of structures observed that connect the two donor layers as shown in the snapshot of Figure 6.7b: the bridge has a constant interfacial area regardless of the thickness of the domain being bridged (which is the main feature being affected by composition). This nearly constant interfacial area translates into comparable energetic penalties associated with a new domain formation, and as a result, a largely unaffected γ. In our previous work[1] we tested similar ABA LCEs where ε^{ij} was below the crystallization threshold that we see here, but we still observed a compositional dependence of ⍰⍰ In that study, two key differences were that (i) all deformations were conducted at constant total volume at a higher density (ρ=0.9), and (ii) the systems were 8 times smaller than those used here. While crystallization was not detected in those simulations, these differences could have induced some change in the deformation behavior which led to the observed compositional dependence.

To quantify the difference between the mechanical properties of the crystalline and non-crystalline domains, we arranged homopolymer chains (20-mers with ε^{ij}=1.0) into the crystalline structure, equilibrated in the NPT ensemble using the same P^*_{set} as that for a

network with f_A = 0.5 and ε^{ij}=1.0. This system was then isothermically deformed along the strain direction, while barostating the fluctuations of box along the perpendicular directions. A similar tensile experiment was conducted with a non-crystalline domain by using a homopolymer T-LCE which is prepared by compressing the network to ρ=0.85, followed by the same equilibration method used for DBC T-LCE where the box dimension barostated independently is the one with which chains primarily align. The calculated Young's modulus of the homopolymer crystalline network, $Y=\Delta\sigma^*/\Delta\alpha$ is 327, while for the homopolymer non-crystalline T-LCE is $Y\sim0.017$. As expected (and observed in experiments[35]) the crystalline domain requires much more energy to be deformed, which translates into a greater toughness. This difference explains the superlinear increase in γ when going from largely amorphous systems ($\varepsilon^{ij}\leq 0.5$) to crystallizable systems ($\varepsilon^{ij} \geq 0.75$) as observed in Figure 6.3 and in Figure 6.12 of the SI.

6.4 CONCLUSIONS

Expanding on our previous work with T-LCE, we have explored the effect of different degrees of block immiscibility (ε^{ij}) on the tensile response. As a preliminary analysis, we show that while block microphase segregation is a dominant factor in the magnitude of the toughness of our BCP T-LCEs, it is only one of many factors that gives rise to the unique saw-tooth tensile behavior. Indeed, such a tensile response results from the concerted effect of three design elements: backbone stiffness, BCP chain composition, and network connectivity. Upon macroscopic uniaxial deformation of the network, chain stiffness results in microstructures that favor local deformations, BCP chains provide an energetic driving force that favors the concerted formation of lamellar domains, and the network connectivity

restricts lamellae from tilting rather than undergoing new domain formation in preserving the lamella spacing.

Previous work postulated that γ was linearly dependent upon ε^{ij} and had a parabolic dependence on composition. In the systems studied in this work, we observe that the proposed compositional dependence is approached but only for $\varepsilon^{ij} \geq 0.75$, while there is no clear compositional dependence for lower values of ε^{ij}. The emergence of these crystalline domains for networks with $\varepsilon^{ij} \geq 0.75$ changes the mechanism of how new lamellar domains form, going from transitional structures resembling a bridge between the two donor layers, to protrusions bulging from each donor layer. Because the interfacial area associated with these protrusions correlates with the volume fraction of the minority block, the toughness in these crystallizing systems exhibits a relationship with chain composition that is different from that postulated in our previous work. The emergence of crystalline domains also increases γ by virtue of their greater Young's modulus compared to that of the non-crystalline domains. The emergence of crystallinity leads to a non-linear relationship between ε^{ij} and γ, which contrasts the linear relationship which would emerge had surface energy alone been the dominant contributor to the tensile response.

The emergence of the crystalline domains was found to be a result of the increase of like-pair interactions (lower "effective" temperatures), high strain, and the increase in interfacial surface tension between blocks. Making all pair interactions in the system the same (as in homopolymer LCEs) removes the interfacial surface tension, and results in more mobile crosslinks that make crystalline domains more difficult to form (requiring higher strains) and to stabilize.

The tensile responses of BCP LCEs under deformations other than uniaxial extension is also of interest for practical applications like uniaxial compression. If a system is compressed in the direction of initial lamellar alignment (z, for example), then the other two dimensions (x and y) would be undergoing extension, which could lead to soft deformations (due to goldstone modes) or other types of domain-transformation deformation behavior. In terms of molecular design parameter space, an avenue of interest would be the introduction of non-constant stiffness across the length of the chains. While chain alignment is needed for the emergence of the saw-tooth tensile behavior, it is unclear whether it can be achieved when different blocks (along the main chain or as side chains) have differing backbone stiffnesses. This unexplored design variable can be leveraged to blend the entropic elasticity from more flexible segments, with the tensile response associated with by the saw-tooth behavior.

6.5 Acknowledgements

This work was supported by the National Science Foundation Award CMMI 1435852. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1053575. The authors are thankful to Abhishek Sharma for his help with the calculations for identifying crystal domains based on a translational order parameter.

6.6 Supplemental Information

6.6.1 Role of Block Immiscibility on Uniaxial Deformation of Non-Network BCPs

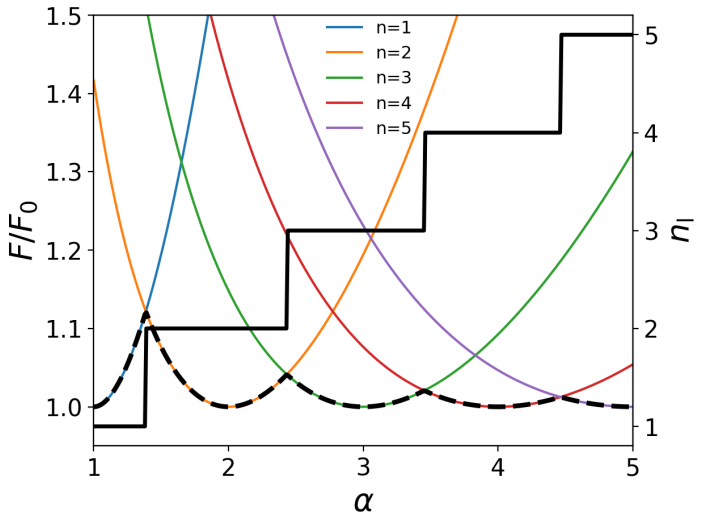

Figure 6.9: Plot of the free energy of a given number of A-B lamellae bilayers, n_l, as a function of the plate separation, converted into extension ratio, α. The dotted black line traces the manifold of minimum free energy which illustrates the change in n_l as the system is deformed. The solid black line traces the value of n_l, corresponding to the minimum free energy at each given value of α.

To test the viability of the behavior suggested by Figure 6.9, a set of simulations was conducted where the initial starting lamellar configurations of BCP T-LCE networks of selected compositions and ε^{ij} values had three bonds to each tetra-functional cross-link removed to reduce the networks to systems of uncrosslinked BCP chains or melts. Subsequently, two deformations simulations were performed. In the first, K_{bend} is the same as in the original T-LCE systems (semi-flexible chains), while in the second simulation K_{bend} is set to 0 (flexible chains). These deformation simulations were performed at a deformation rate of $8\times10^{-6}\,\alpha/\tau$ for computational efficiency. For the semi-flexible chains the deformation methodology is identical to the isobaric conditions of the full network, while for the flexible

chains the deformation is performed at constant volume to maintain a consistent density across systems, given that the equilibrium pressure will change when K_{bend} is set to 0.

Figure 6.10 shows a representative set of curves for semi-flexible and flexible chain melts. The semi-flexible system shows a single stress peak corresponding to the formation of a new single domain, followed by a monotonic stress increase. The absence of subsequent stress peaks is due to the lamellae tilting in the box to keep the domain spacing constant, rather than forming new domains (Figure 6.10). The flexible system shows this lamellar tilting over the initial range of deformation but a stress peak appears at $\alpha \backsim 3.25$ where this tilting can no longer be accommodated in the relatively small lateral box dimension, and the system undergoes a transition involving a lamellar reorientation coupled with a new lamellar layer formation. While the periodic boundary conditions in our system do not perfectly represent a bulk system, they provide a sensible approximation as can be seen in Figure 6.10 where the interconnection of the segregated domains across the boundaries emulates a large domain despite the small box lateral dimension at large α.

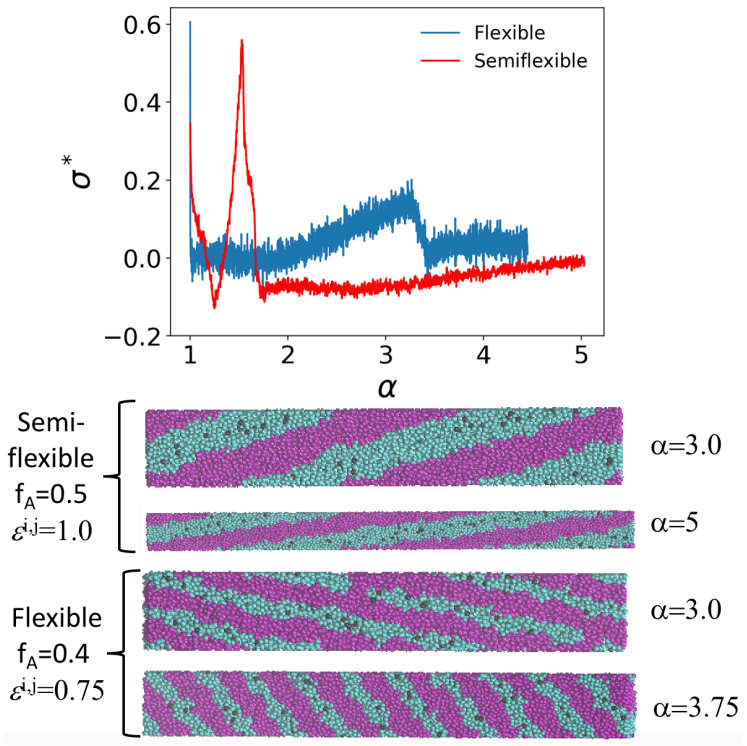

Figure 6.10: *Top:* Example tensile responses for systems which are generated by removing the 3 of the 4 bonds to each cross-link, effectively uncrosslinking the network into a melt of linear triblock chains (Semi-flexible chains: f_A = 0.5, ε^{ij}=1.0, Flexible chains: f_A = 0.4, ε^{ij}=0.75). The semi-flexible system has the same bending constant as the original network, while the flexible system has zero bending constant. *Bottom:* Snapshots from two tensile responses show in the plot.

6.6.2 Values of P^*_{set}

Table 6.2: P^*_{set} used for all tested systems

System	ε^{ij}=0.25	ε^{ij}=0.5	ε^{ij}=0.75	ε^{ij}=1.0
20A	3.85	3.75	3.20	2.4
30A	3.95	3.9	3.35	2.65
40A	3.95	3.85	3.2	2.75
50A	3.95	3.9	3.45	2.78
60A	3.95	3.8	3.35	2.65
70A	3.94	3.85	3.32	2.6
80A	3.81	3.66	3.1	2.4

6.6.3 Surface area values

The surface area is calculated by first calculating the structure factor of the minority phase for snapshots from the deformation of each system, inverting the structure factor onto a rectangular grid, and applying a cutoff such that if the density at a given grid point is above that density, it is considered to be a point the minority phase (considered the majority phase otherwise). For each minority phase grid point, if any of the 6 neighboring grid points are the majority phase, the rectangular area shared between these two grid points is added to the total surface area. For systems where $\varepsilon^{ij} \leq 0.5$, the average surface area is calculated, and shifted such that the interfacial area is zero at $\alpha=1$.

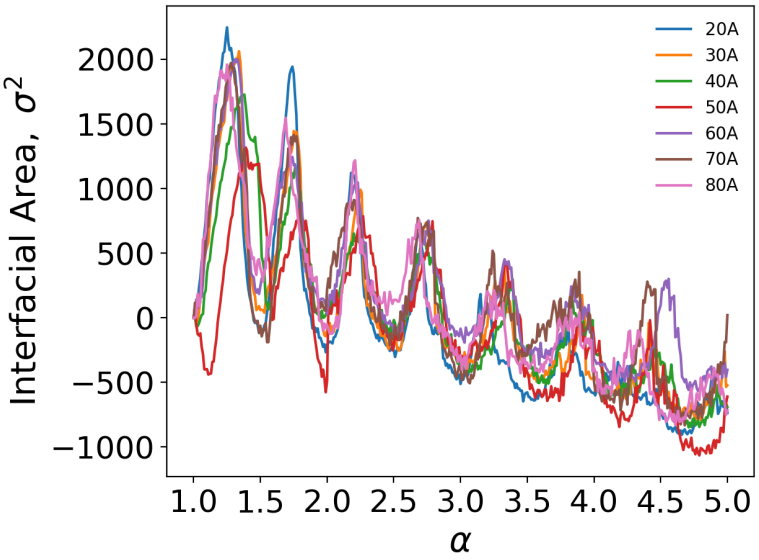

Figure 6.11: Plot of the average interfacial Surface area for systems with $\varepsilon^{ij} \leq 0.5$. The curves are shifted such that the interfacial area is zero at α=1, because we are interested in the change in the surface area, not the absolute value.

6.6.4 Values of MPDL

To quantify the difference in deformation mechanism between crystalline, and non-crystalline networks, the minimum probability density of the minority phase between the donor layers (MPDL) is calculated. The average value of MPDL for all compositions and values of ε^{ij} is presented in Table 6.3 where the value is normalized to the value of MPDL observed at the given composition, and $\varepsilon^{ij}=0.25$. The average values of MPDL are given in the last row, showing the distinct change in value when $\varepsilon^{ij} \geq 0.75$.

Table 6.3: Average values of MPDL for all systems tested, normalized to the value at $\varepsilon^{ij}=0.25$ for a given composition.

System	$\varepsilon^{ij}=0.25$	$\varepsilon^{ij}=0.5$	$\varepsilon^{ij}=0.75$	$\varepsilon^{ij}=1.0$
20A	1	0.5884	0.5026	0.4664

30A	1	0.8075	0.3785	0.3563
40A	1	0.8724	0.3998	0.2169
50A	1	1.0335	0.5484	0.4874
60A	1	1.4127	0.4680	0.2748
70A	1	0.9531	0.1895	0.2016
80A	1	0.7971	0.2082	0.3701
Average	1	1.007	0.4215	0.3195

6.6.5 Average Toughness Values

Table 6.4: Unnormalized average toughness values for all systems, and compositions tested

System	ε^{ij}=0.25	ε^{ij}=0.5	ε^{ij}=0.75	ε^{ij}=1.0
20A	1.308	2.238	3.311	4.808
30A	1.433	2.209	4.020	5.764
40A	1.543	2.019	4.154	6.238
50A	1.545	2.384	4.662	6.878
60A	1.649	2.398	4.367	6.418
70A	1.480	2.287	3.962	5.674
80A	1.583	2.212	3.465	4.846

The influence of the crystalline domains observed for the $\varepsilon^{ij} \geq 0.75$ T-LCEs can be seen in Figure 6.12 where γ is plotted against ε^{ij}, where the relationship between γ and ε^{ij} shows a non-linear trend which can be tentatively fitted to two linear regimes intersecting at $\varepsilon^{ij} \approx 0.5$. Fitting a straight line between the values of γ at f_A = 0.5 for non-crystalline networks (ε^{ij}=

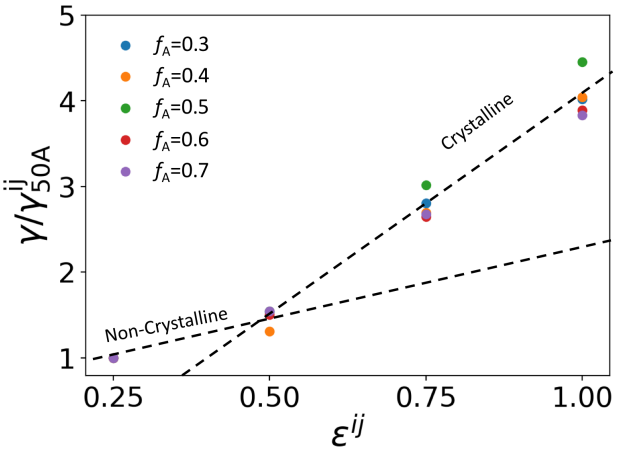

Figure 6.12: Plot of γ as a function of ε^{ij} for all compositions tested. The dashed lines are tentative extrapolations of the non-crystalline and crystalline regimes. f_A = 0.2 and 0.8 systems are omitted for clarity.

0.25, 0.5, non-crystalline line in Figure 6.12) and extrapolating to ε^{ij}=1 assuming no crystallization) gives $\gamma \approx 4.05$ (for $f_A = 0.5$) a value that would align with the expectation that γ should be proportional to ε^{ij}. The values for $f_A = 0.2$ and 0.8 are omitted from the graph for clarity because they deviate from the other systems. However, such a linearly extrapolated value is significantly smaller than the calculated value of 6.88 given in Table 6.4, indicating that the crystalline regions generate a greater resistance to deformation that is not fully accounted for by the changes in interfacial energy with ε^{ij}, but that could be attributed to having a greater Young's modulus, Y.

6.6.6 Dependence of Composition on the Surface Area Created During New Domain Formation

In this model, a system of 20-mer chains, identical to the network chains in the T-LCE, are arranged on a cubic grid with spacing $2^{1/6}\sigma$ and periodic boundaries (example snapshot shown in Figure 6.13a). A single chain is displaced in increments of $2^{1/6}\sigma_{LJ}$ (Figure 6.13b) along the director, and the total non-bonded energy is calculated using the model employed in our tensile experiments. This calculation is repeated for several chain compositions. For each composition the values of the non-bonded energy are shifted to zero when the chain displacement is 0. The energies are normalized with respect to the maximum observed energy (to give U_{int}) and the displacement is normalized with respect to the chain contour length (to give the Chain Displacement Fraction). These quantities are plotted in Figure 6.13c which shows that for all compositions a plateau energy is reached which is nearly proportional to volume of the minority block, $U_{int} \propto \min(1\text{-}f_A, f_A)$.

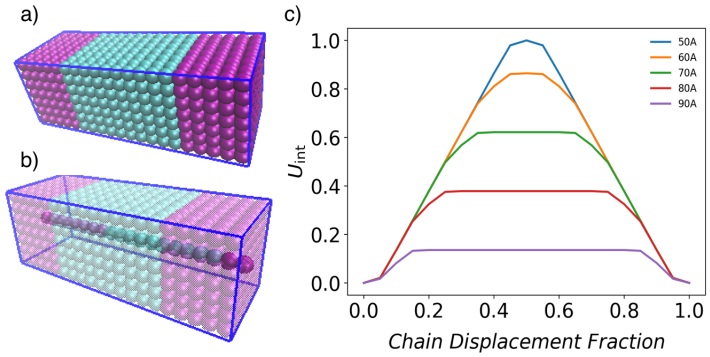

Figure 6.13: a) Snapshot of a model system of 20-mer ABA triblocks arranged on a cubic lattice to give flat interfaces between the A and B domains. b) The same system with a single chain displaced by some amount along the strain axis. This displacement causes an increase in the non-bonded energy, which is calculated for the entire range of displacements. c) A plot of the normalized interfacial energy U_{int} as a function of chain displacement for different block compositions.

7 Accelerated discovery of atomistic chemistry from coarse-grained simulations

7.1 Introduction

The ever-expanding availability of computational resources has fueled a fast growth in the size and scope of the molecular simulations currently used for property prediction. Indeed, with resources such as XSEDE, simulations involving hundreds of processors, and millions of atoms are potentially viable[59,107,213,229]. Despite these advances, there still exist many physical and chemical processes whose length scales are beyond the reach of computationally accessible timescales, such as those involving large biomolecules and other macromolecules. Indeed, such simulations often encounter rugged free-energy landscapes and kinetic trapping in deep metastable basins. To address these kinetic barriers, techniques such as Parallel Tempering/Replica Exchange[51,170,198,203], Metadynamics[12,13,120], Pransition-path Sampling[42,54], Kinetic Monte Carlo[15,26], and Time-Temperature Superposition[134] have been developed and used.

One of the most successful approaches to speed up molecular simulations has been the use of coarse-grained (CG) models, i.e., models that bundle groups of beads from a more detailed model into single beads to thus eliminate microscopic degrees of freedom that are not essential to resolve structural details above a certain length scale. CG models have smoother potential energy surfaces that are easier to sample ergodically compared to their all-atom (AA) counterparts, whose rougher potential energy landscape can create kinetic traps[92].

The CG models for macromolecules can be broadly classified into two categories: (i) “chemical” models if mapped directly from a Chemistry-specific (CS) polymer and (ii) “physical” models if intended to describe a broad class of polymers. In the former directly-

mapped CG or "DCG" models, their parameters are calibrated to match selected results of properties obtained from experiments or AA simulations of the material of interest. In the latter case, one begins with a relatively small simulation of the AA molecule and a CG molecule with a specific CG mapping (i.e., a recipe for the way how atoms in the AA molecule are mapped into the different CG beads); such a mapping is often guided by physical intuition. Once suitable functional forms have been selected for the bonded and non-bonded interaction potentials, the model can be parameterized by such methods as iterative Boltzmann inversion[136,169], force matching[53,93], or relative entropy[25,187]. The goal of the parameterization is to construct CG molecules such that their behavior mimics that of the known molecule at a prescribed the level of detail. The degree of coarsening (DOC); i.e., how much detail is averaged out, can range from simply integrating out the hydrogens like in typical United Atom (UA) models[49], to lumping entire monomers/amino-acids or even long sections of polymer coils into single beads as in models originally used with Dissipative Particle Dynamics (DPD). Due to the simplicity and computational efficiency of CG models with greater DOC, many studies have adopted generic, physical CG models intended to capture the typical behavior of a class of polymers rather than that of any specific polymer chemistry. While many of these studies are able to reproduce experimentally known physical trends of some macromolecules, others can also reveal new or unusual results. In the latter case, it would be of interest to identify specific molecular chemistries (i.e., AA models) that could be good candidates to capture the generic CG molecules investigated. The process of determining the identity of these AA molecules is henceforth referred to as inverse coarse-graining (ICG).

Figure 7.1 schematically compares DCG and ICG. If it is known that a given AA molecular model has a specific property of interest, then DCG would be a suitable approach to explore perturbations in behavior in a close proximity of compositional space. If the goal is to widely explore a potentially novel type of behavior, ICG would be a suitable methodology, because generic CG molecular modes are typically "coarser" and more computationally efficient. Once a CG model has generated results of interest, candidate AA CS-molecules need to be determined to guide experimental efforts towards realizing such predictions. ICG can thus become a powerful strategy in materials design, well aligned with the objectives of the Materials Genome Initiative[39,94]. However, ICG has not been as well studied as DCG, partly due to the ill-posed nature of the ICG problem compared to DCG. Indeed, for a given AA CS-molecule and CG model, there exists in principle one optimal set of model parameters, but for a given CG model, many different AA CS-molecules can be mapped onto the same CG molecule. As such, a directed evolutionary approach would be highly desirable, so that the candidates for the optimal AA CS-molecule can be evaluated and evolved towards the best fit of the CG model, ideally, in an automated way (e.g., aided by machine-learning techniques) [7,27,58,115,224].

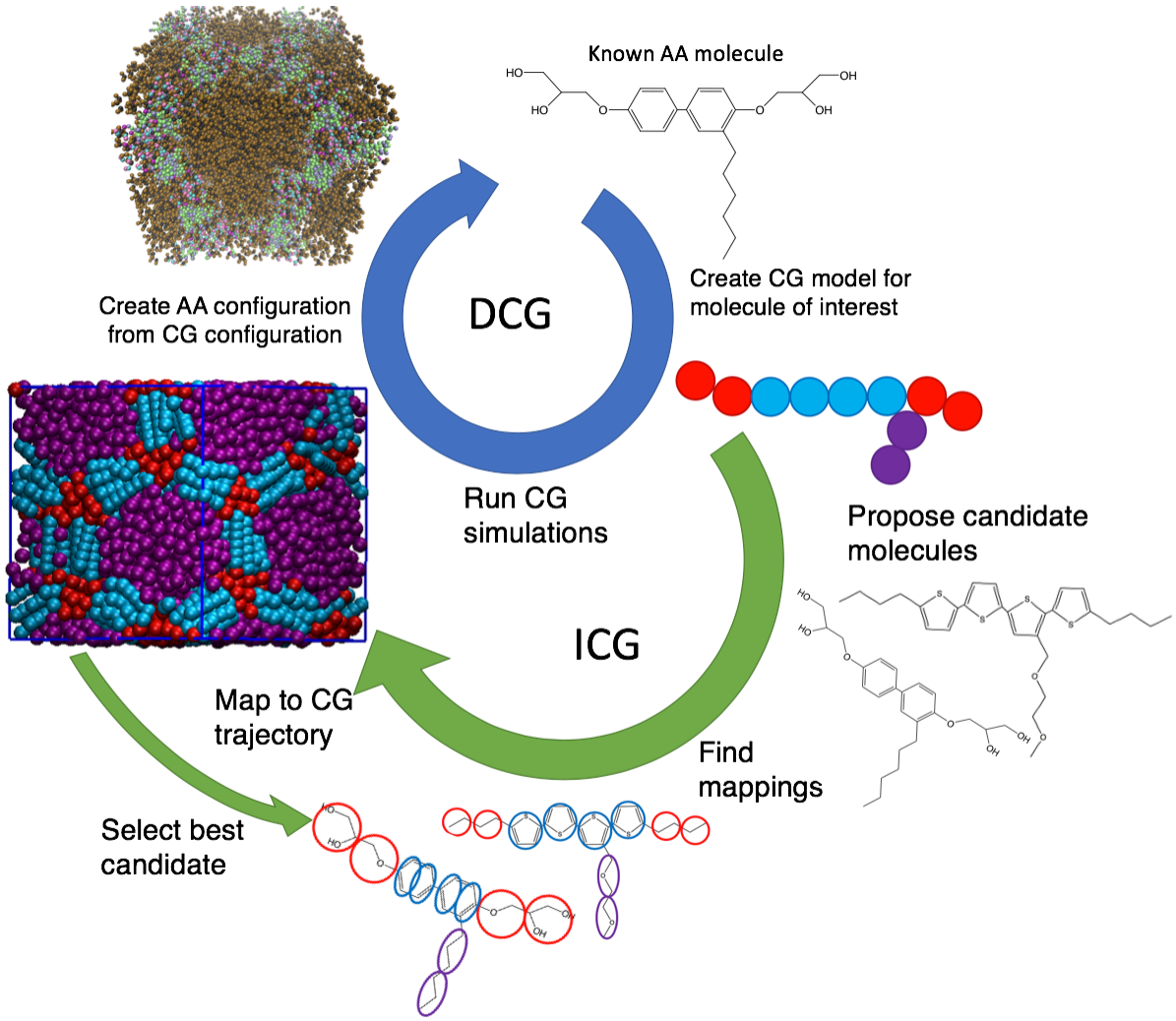

Figure 7.1: Schematic of the DCG (blue arrow) and ICG (green arrows) processes. DCG begins with parameterizing a CG model based on simulations of the original AA CS-molecule. The CG trajectory can then undergo Reverse Coarse Graining (RCG) to return an AA structure. In ICG, only the CG model is known and candidate molecules are proposed. Each CS-molecule has an optimal mapping that most closely replicates the original CG model. The trajectory of each CS-molecule is coarse-grained using this mapping into a trajectory of the CG model. These trajectories are compared to yield a best candidate.

In this work we propose a flexible methodology for determining which AA CS-molecule from a given pool of candidate molecules is the best fit for a target CG model based on minimizing an objective function. Our approach is similar to that used in the relative entropy (RE) methodology[116]. Shell et al.[25] used the concept of RE to iteratively parameterize a CG

model in DCG, where at each iteration a system of the CG molecules is equilibrated under a given set of potential parameters (defining the CG Hamiltonian $\mathcal{H}_{\mathrm{CG}}$) to get the "true" CG trajectory, T_{t}. Then a separate equilibrated trajectory of the AA CS-molecular system is mapped to return a new CG trajectory, T_{AA}. $\mathcal{H}_{\mathrm{CG}}$, and a function f which depends on $\mathcal{H}_{\mathrm{CG}}$, are evaluated for each frame of T_{t} and T_{AA}, effectively "simulating" both trajectories under the CG forcefield. The objective function, Φ, measures the degree to which T_{t} and T_{AA} differ by evaluating $\Phi = f_t - f_{\mathrm{AA}}$, where f_i is the average value of the function f obtained when "simulating" T_{i}. This difference determines the changes to be made to $\mathcal{H}_{\mathrm{CG}}$ (i.e., the CG model parameters). This process is repeated until $\mathcal{H}_{\mathrm{CG}}$ converges and hence Φ is minimized.

In our methodology for ICG reverse mapping, $\mathcal{H}_{\mathrm{CG}}$ is fixed (i.e., the CG model doesn't change), and what changes is the candidate CS-molecule being considered. This can also be thought of as changing $\mathcal{H}_{\mathrm{AA}}$. For this purpose, a system for each candidate CS-molecule is simulated and coarse-grained according to a mapping which best satisfies the constraints of a desired CG molecule to give T^i_{AA}. The procedure to generate these mappings is described below. At this point we can calculate f^i_{AA} and f_{CG} for all T^i_{AA} and T_{t}, respectively, where the candidate CS-molecules which minimize

$$\phi = abs\left(f_{CG} - f^i_{AA}\right) \quad (7.1)$$

will be kept for further application of a machine learning algorithm to propose new candidate CS-molecules. A schematic comparison of our method to the original RE method is given in Figure 7.2.

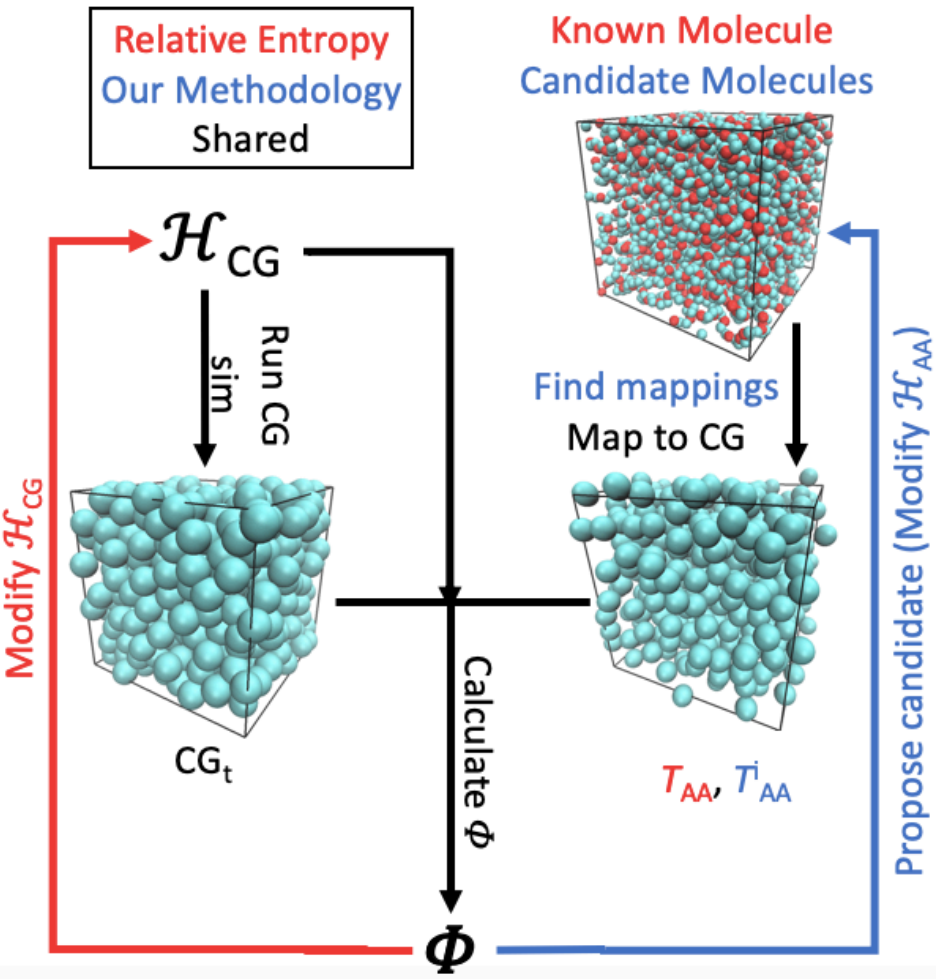

Figure 7.2: Schematic comparison between the relative entropy (RE) framework and the proposed methodology, where items belonging to just RE, just ICG, or both are colored in red, blue, and black, respectively. With a given $\mathcal{H}_{CG}$, a CG simulation is run to give a "true" trajectory of the CG model, T_t. In RE the AA CS-molecule is known, while in ICG multiple CS candidates are proposed and the trajectory of each is mapped onto the CG model to give T_{AA}. In RE this mapping is known a priori, while in ICG the optimal mappings must be found. Once mapped, T_t and T_{AA} are used to calculate the objective function Φ. In RE the Φ values are used to modify $\mathcal{H}_{CG}$, while in ICG they are used to identify the best candidate CS-molecules, which can in turn be used to propose new candidates.

Before different AA CS-molecules can be compared, the way in which the AA atoms are partitioned (mapped) to the CG molecule must be determined. As stated previously, mappings must satisfy the constraints of the CG model (such as number of beads and bond topology). However, there can exist multiple mappings that satisfy the constraints and

several studies have explored how different mappings affect the ability of the CG molecule to reproduce properties of the AA CS-molecule[80,102]. In this study, the mapping needs to best reproduce the properties of the CG molecule, not vice-versa. Every mapping, "s", of a given CS-molecule onto the target CG model will return a different value of the objective function, Φ_s, so only the "optimal" mapping which minimizes this value should be used when comparing across candidate AA CS-molecules. To facilitate finding this optimal mapping, a methodology to automate the generation of initial mappings is proposed. For each candidate CS-molecule, a process similar to ICG is followed except that, we now know the CS-molecule and CG model so what is modified through the algorithm is the mapping itself, until Φ is no longer minimized.

For the automatic generation of a CG molecule from a given AA CS-molecule, an existing approach is exthe CGTools plugin for VMD[87] wherein a neural network learns to map the AA molecule based on its "shape", i.e., the position and connectivity of the constituent atoms. Coarse-graining by the "shape" is also the basis of a dimensionality reduction method[118,122] which uses data graphs and diffusion maps, where in the context of coarse-graining our AA CS-molecule, the graph is the AA CS-molecule and the dimensionality reduction is the coarse-graining process. While these two approaches show promise, neither one operates under the key constraint that ICG imposes: that the final CG molecule is known while the CG mapping is not.

Even for the more well-posed problem of DCG, the process of parameterizing a CG model is laborious and software like VOTCA[174] has been developed to automate the process. We propose an equivalent tool kit that is focused on making the ICG process more automatic.

The main capabilities of this toolkit are: (i) finding an appropriate mapping for the AA CS-molecule (useful for both DCG and ICG), (ii) determining a one-to-one correspondence between the beads of a proposed mapping and the beads of the target CG molecule, and (iii) finding repeating motifs in a given molecule through a compression algorithm. These capabilities are used throughout the process of finding the best AA CS-molecule for a CG model, but can also be used in other contexts, such as in the RCG[67,145,176,220] process that maps atomistic detail onto the CG model making our toolkit have broader uses as a molecular simulation toolkit.

The rest of the manuscript is organized as follows. We describe our simulation model and the methods for proposing new mappings, for selecting which optimal mappings for each candidate CS-molecule, and for comparing among different candidates. The following section describes the implementation of our methodology for finding the optimal CS-molecules to different CG models with varying DOC. We conclude by assessing the performance of our methodology and look at future avenues for improvement.

7.2 Models & Methods

7.2.1 Simulation Models

A broad range of forcefields were used which include the OPLS-AA/UA[99], MARTINI[98], DPD[56,74], KG bead-spring model[72], and new forcefields derived for specific molecules using DCG[75,195]. For simulations using unscaled units, thermo/barostating was done using the Nosé-Hoover Thermostat/Barostat to maintain a temperature of 300K and a pressure of 1atm, with timesteps of 1fs. For simulations using scaled (Lennard Jones) units, simulation were run in the NVT ensemble using the Nosé-Hoover Thermostat at $T^*=1$,

ρ^*=0.85, and timesteps of 0.005τ. In all cases, a melt state is simulated by using a varying number of molecules depending on how large/small the molecules are. Detailed information about our simulation systems, parameters and methodologies are given in the SI (section 6.6).

7.2.2 Finding Mappings

While numerous mappings of the AA CS-molecule onto the CG model may exist, many of them can lead to stretched bonds, or many more beads being mapped to one bead than another even if they are the same type in the CG model leading to poor representation of the CG model. As such, an automated process to generate mappings for evaluation is required. The proposed code requires two inputs, an AA information file containing one AA molecule and a similar file for the CG molecule. These files need to contain information regarding atom positions and types, as well as the bond structure of the molecules. An overview of the algorithm is presented in Figure 7.3. Importantly, although our main focus is on going from an atomistic level of detail (AA/UA models) to a CG model, this approach can be used to find mappings between any two models that differ in the level of atomic description.

The algorithm begins by finding η_{CG}, the ratio of total number of heavy atoms (non-hydrogen) to the number of desired CG beads. When rounded down, η_{CG} gives an average "size" of each CG bead, $\langle\eta\rangle$. By disregarding hydrogens, any atomic level description is reduced to the UA representation so the input model of the CS-molecule can be either AA or UA. In the next step, terminal atoms are assigned to their own groups (black circled atoms in Figure 7.3). A terminal atom is defined as one which is only bound to 1 other heavy atom which hasn't been

assigned to a group yet (herein referred to as an "unmapped" atom). Because $\langle\eta\rangle$ is usually greater than 1, groups need to be "grown" to $\langle\eta\rangle$.

To grow the group with initial atom i, g_i, the "batch" of atoms which can be added to g_i is found. A batch is defined as all unmapped atoms, j, with a given bond separation number (the minimum number of bonds separating i and j, BS_{ij},). For example, the first (second) batch is all unmapped atoms with BS_{ij}=1(2). Batches are indicated by the atoms inside the dashed oval of the same color in Figure 7.3. A batch is only calculated for "active" groups having $\eta_i<\langle\eta\rangle$. Active groups are marked by non-black solid circles/ovals in Figure 7.3. Groups with no atoms in the current batch or with $\eta_i=\langle\eta\rangle$ are labeled as "finished" and are no longer grown. Finished groups are marked by the non-black rectangles in Figure 7.3. All atoms in a batch are added to g_i of any active group if $(\eta_i+\eta_{batch})\leq\langle\eta\rangle$ and no atom in the batch is part of another batch.

When adding a batch to an active group, inter-batch and intra-batch conflicts can arise. In the former case, an atom in the batch for g_i is also in the batch of at least one other active group, g_j. The strategy to circumvent this is to evaluate each way these conflicts can be resolved (i.e., the atom shared between two or more batches is assigned to only one of the groups). If N batches share the same atom, then there are N ways to resolve the conflict. For the first such resolution, the atom is assigned to the appropriate group. However, to evaluate all N ways to resolve the conflict the algorithm must be recursively called N-1 times, each time carrying over information about the mapping, finished groups and how the conflict was resolved. This is illustrated in Figure 7.3 where the red and green groups have an inter-batch

conflict, which is resolved by assigning the shared atom to either group resulting in the original instance of the algorithm continuing and one recursive call of the algorithm.

In the case of intra-batch conflicts, the size of the group, η_i, plus the size of the batch exceeds $\langle\eta\rangle$. Similar to inter-batch conflicts, this is addressed by only adding a subset of the batch atoms such that, $\eta_i+\eta_{subset}=\langle\eta\rangle$, where permutations of the subset initiate a new instance of the algorithm (similar to inter-batch conflicts). Once inter- and intra-batch conflicts have been resolved, the presence of any remaining active groups is checked. All active groups are grown simultaneously until no active groups remain, at which point any remaining unmapped heavy atoms are processed by creating and batch-wise growing new terminal groups. This cycle continues until all atoms are mapped, upon which the algorithm ends and the mapping is reported. A simple molecule is used in Figure 7.3 to illustrate the process, with an example of how the algorithm generates initial schemes for more complex chemistries in the Supplemental Information (SI, section 7.6.1), based on our previous work[46].

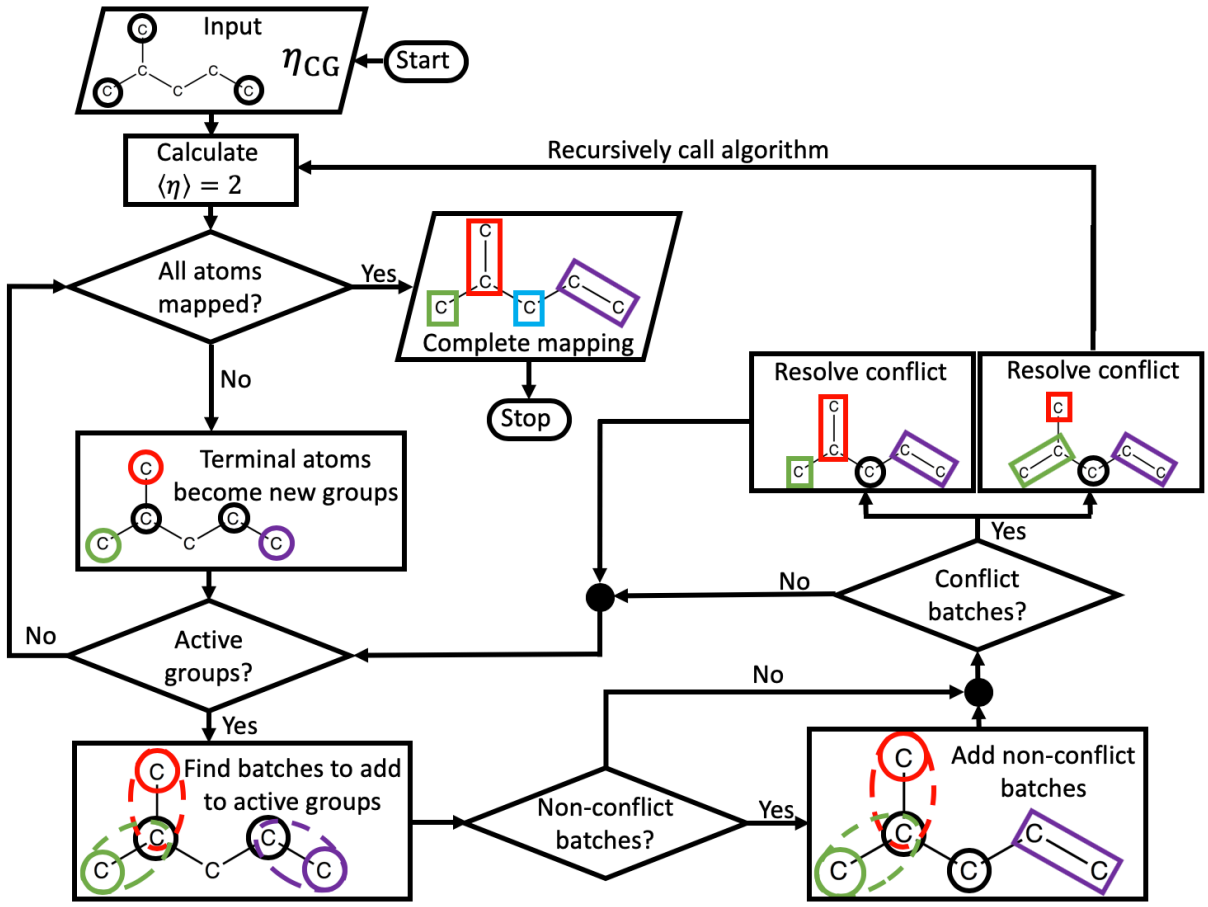

Figure 7.3: Depiction of algorithm to automatically generate CG mappings for a candidate molecule. It begins with a) stripping the molecule of hydrogens to give the UA representation, and placing the first initial groups starting at the terminal atoms. b) Groups are grown until $\eta_i \leq \langle\eta\rangle$; however, if there is an atom that has been assigned to two or more groups, then new algorithm calls are initiated for each permutation where the conflicting atom is assigned to one of the groups. c) Once all "active" groups are grown, atoms which are bonded to an atom already assigned to a group are labeled as terminal atoms and the algorithm is iterated.

By design this algorithm will always give a number of groups equal to or greater than n_{CG}. Thus, some groups may need to be merged so that the number of groups is equal to n_{CG}. Potential mergers are identified by pairing the smallest group(s) in the mapping with each of their smallest neighboring group(s). Each such a possible merger is tested by again recursively calling an algorithm similar to that used for finding the initial mappings. This process continues until the number of groups is equal to n_{CG}. This overall strategy of

proposing many new mappings increases the likelihood of finding the optimal mapping because even a single misassigned atom may cause a mismatch with the bond topology of the desired CG molecule. While this algorithm for generating initial mappings is intended for automating ICG, it can also be helpful with DCG for objectively searching multiple feasible mappings, some of which may not have otherwise been considered by the researcher.

7.2.3 Finding repeating motifs

The method described above is suitable for molecules where it is computationally manageable to find all non-conflicting mappings. However, this is not the case for large macromolecules that would engender an intractably large number of initial mappings. Since these macromolecular systems often contain many repeat units, identifying them would greatly reduce the combinatorial redundancies as changes to the mapping within one repeat unit could then be propagated to all repeat units. Indeed, finding repeat units reduces the problem of finding the mapping of a macromolecule to the tractable problem of finding a mapping for a small molecular repeat unit. Identifying repeat units in a small molecule can also be helpful to speed up the process of enumerating possible mappings.

To identify these repeat units, we employ the simplified molecular-input line-entry system known as SMILES[214,221], a methodology for representing the topology of a molecule as a linear string, called the "smile", where repeat units show up through a recurring pattern in the string. Our procedure for creating a smile for a given molecule follows the standard procedure, by tracing the "backbone" of the molecule determined as the path connecting the atoms with the maximum value of BS_{ij}. Due to the way the smile is constructed, each molecule

has two smiles associated with it, created by starting at either end of the "backbone". One difference between the typical way a smile is created and our methodology is that we do not break up any ring but instead replace it by a "superatom". This modification avoids complications with how to open rings, and aligns with established CG techniques where rings are generally treated as single beads. With each smile, a string compression method is applied, similar to that used in the zip file format[236,237] as detailed in the SI (section 7.6.2).

7.2.4 Automated/Machine Learning Mapping

Once a set of initial mappings with the correct number of groups has been created, an algorithm is used to ensure that the bond topology of a mapping matches that of the desired CG molecule. The algorithm begins with the calculation of BS_{ij} and the termination map TM_{ij} for both the mapping under evaluation (SBS_{ij} and STM_{ij}) and the desired CG molecule (DBS_{ij} and DTM_{ij}), respectively. The termination map is a metric of the network connectivity where for a given atom i in the molecule, TM_{ij} is the number of atoms, k, such that BS_{ik}=j are also bonded to at least two heavy atoms.

We also create the permutation matrix, P, which is a binary matrix describing if atom i in the mapping and atom j in the desired CG molecule can (P_{ij}=1) or cannot (P_{ij}=0) be assigned to each other. A complementary correspondence matrix, C, is also created whose initial default entries are 0. A "correspondence" occurs when only an atom in the mapping, i, can be represented by a unique atom in the CG molecule, j:

$$\text{if } \sum_j P_{ij} = 1 \text{ and } P_{ij} = 1 \Rightarrow C_{ij} = 1 \quad (7.2)$$

A complete correspondence of beads has been found if there is one and only one entry of 1 in each row, and column in P (all other entries being zero).

At this point several consistency checks are performed. The first check involves examining whether for each pair of i and j with P_{ij}=1 that the termination map of both i and j are the same:

$$P_{ij} = \begin{cases} 1 \text{ if } STM_{ik} = DTM_{jk} \forall k \\ 0 \text{ otherwise} \end{cases} \quad (7.3)$$

Then if C_{ij}=1 a bond separation distance check is done:

$$\text{if } C_{ij} = 1 \text{ and } SBS_{ik} \neq DBS_{jl} \Rightarrow P_{lk} = 0 \quad (7.4)$$

This check is repeated until no more entries change in the permutation matrix. After these two checks, the only remaining atoms that can have more than one correspondence are symmetric atoms, i.e., atoms of the same type with the same neighboring bond topology. Examples of symmetric atoms in an AA model would be the hydrogens of a methyl group and ortho-position carbons, and in CG models would be the two end beads on a linear homopolymer chain. A decision must be made to break the symmetry and make the correspondence unique. After identifying a symmetric atom i, for each atom j such that P_{ij}=1, the algorithm is called recursively, where the input is the permutation matrix where atom i only corresponds to atom j, and then the bond separation is re-checked. This process of selecting the correspondences of the symmetric atoms is repeated until no more symmetric atoms remain. Finally, if an atom i that has no possible correspondences (i.e., $\sum_k C_{ik} = 0$), the mapping in question is discarded as it cannot map onto the desired CG molecule.

7.2.5 Proposing new mappings for minimizing Φ

Once a set of initial mappings to the desired CG model have been found, we then proceed to optimize them by introducing sequential modifications. This is done by "shifting" the boundaries between two groups: If an atom in group i is bonded to an atom in group j, then that atom (and any terminal atoms bonded to it that are also in group i) becomes part of group j. Once the change has been made, it is checked that the topology of the CG molecule still matches the desired one, and that all groups are contiguous.

7.2.6 The Objective function

Regardless of the forcefield used for the AA and CG systems, the goal is to match the behavior of the AA CS-molecule to that of the CG molecule with respect to key geometrical and energetic details of the CG model. This includes matching such observables as the components of the non-bonded and bonded interaction energies. We adopt as a starting point for the objective function f the average potential energy of the system, E_{pot}, because most energy components of $\mathcal{H}_{\text{CG}}$ increase when the system deviates from equilibrium (e.g., arising from stretched bonds/angles), so that the optimal system tends to minimize E_{pot}. Additionally, since beads of the same type in a CG molecule are meant to have nearly identical constituent atoms, they should have similar masses. Hence, an additional factor in the functional form of f is included to penalize large variances in the masses of same-type CG beads (as determined from the coarse-graining of the AA CS molecule):

$$f = E_{\text{pot}} + K \sum_{\text{i}} \frac{\sigma_{\text{i}}}{\langle M_{\text{i}} \rangle} \quad (7.5)$$

where K is a penalty factor in energy units, M_{i} and σ_{i} are the average mass and standard deviation of the mass of CG beads of type i, respectively. The implementation of this penalty

is only applied when searching for the optimized mapping of a given candidate, otherwise, K=0. The value of K will vary based on the CG model, with larger values indicated for models with softer potentials.

The use of scaled units, like Lennard-Jones units, in some CG models creates an ambiguity in their length scale. Because the bond types are known for the bonds in the CG molecule, the average bond length of each of these bonds can be found from the trajectories T_t and T_{AA}. Assuming all bond potentials are roughly harmonic and have approximately similar force constants, the bond scaling, χ, can be found as:

$$\chi = \frac{\sum_i a_i c_i}{\sum_i a_i^2} \qquad (7.6)$$

Where a_i and c_i are the average bond lengths for the bond of type i, calculated from T_{AA} and T_t, respectively. Once χ is calculated, it is used to scale the coordinates for all frames in T_{AA}. Once T_{AA} has been thus processed, $\mathcal{H}_{CG}$ is evaluated for each "simulated" frame to calculate f_{AA} and the same is done for calculating f_{CG} from T_t. For force fields that would only lose a small amount of detail upon coarse-graining (like the UA model), or when searching for the optimal mapping of a given molecule, the previously defined form for Φ in Eq. 7.1 is suitable. However, for CG models losing larger amounts of AA detail, it will be more difficult to discriminate between values of Φ for each candidate. In such cases, a more discriminating form of Φ is formulated which assesses the differences between the radial distribution functions, g(r), from T_t and T_{AA}. A similar approach is used in Iterative Boltzmann Inversion where the difference between the two g(r) functions are used to improve $\mathcal{H}_{CG}$. Here the difference between the two g(r) is quantified using:

$$\Phi = \sum_i^n \sum_{j \geq i}^n \int_{r=0}^{r=r_{cut}} \left| g_{CG}^{ij}(r) - g_{AA}^{ij}(r) \right| \left(\frac{r^*}{r}\right) dr \qquad (7.7)$$

where $g^{ij}_{CG}(r)$ and $g^{ij}_{AA}(r)$ are the g(r) between beads of type i and j for the T_t and T_{AA} trajectories, respectively. The g(r) used here include intra-molecular interactions, as appropriate with the non-bonded interaction scalings for 1-2, 1-3, and 1-4 interactions. To better match g(r) at low values of r, the difference between $g^{ij}_{CG}(r)$ and $g^{ij}_{AA}(r)$ is weighed by the ratio of r^* to r, where r^* is the first peak of $g^{ij}_{CG}(r)$. The integral is evaluated over all non-bonded interaction pairs i,j from r=0 to r=r_{cut} (r_{cut} is the potential cutoff). Again, as with the formulation of Φ for unscaled systems, the chemistry which minimizes Φ is the best fit.

7.3 Results

7.3.1 Small Molecule UA

To validate the proposed method's consistency, we tested that the best candidate CS-molecule for a given UA model is the corresponding AA model. As seen in Table 7.1: Table of Φ values for candidate CS-molecules fitting onto a set of UA models of two different topologies (*Top*: linear, *Bottom*: branched). The best candidate for each UA model is highlighted in green.

	UA Models				
Candidates	**Dimethyl ether**	**Dimethyl sulfide**	**Acetonitrile**	**Propene**	**Propane**
Dimethyl ether	**0.115**	**38.0**	**129**	**2.77**	**2.65**
Dimethyl sulfide	**47.2**	**0.138**	**313**	**77.7**	**21.7**
Acetonitrile	**52.7**	**120**	**0.530**	**46.9**	**60.1**
Propene	**4.08**	**42.3**	**85.7**	**0.114**	**3.92**
Propane	**3.84**	**20.9**	**154**	**10.5**	**0.253**
Dimethylamine	**0.921**	**30.9**	**143**	**4.34**	**1.70**
Perfluorohexane	**611**	**223**	**1270**	**802**	**427**

Hexane	**66.1**	**38.1**	**482**	**63.0**	**63.1**
Ethylamine	**2.80**	**27.0**	**144**	**5.05**	**3.01**
Dichloromethylene	**40.7**	**4.09**	**254**	**64.7**	**16.8**

	UA Models			
Candidates	**Acetone**	**Isobutane**	**Isobutylene**	**DMSO**
Acetone	**1999**	**14276**	**5244**	**63994**
Isobutane	**34853**	**1906**	**15919**	**29826**
Isobutylene	**7348**	**68089**	**135**	**50300**
DMSO	**55151**	**22928**	**48257**	**204**

, two different topologies of UA molecules (linear and star) were simulated and the best fit for each UA molecule from a pool of candidate CS-molecules was found. We include the AA CS-molecule of each of the UA molecules so that the correct fit is known a priori. Indeed, our methodology returns the correct candidate molecule fit for each UA molecule tested. Due to the small degree of coarse-graining entailed by the UA model, the K factor in Eq. 7.5 was set to zero since any disparity in bead masses would lead to large increases in E_{bond}.

Table 7.1: Table of Φ values for candidate CS-molecules fitting onto a set of UA models of two different topologies (*Top*: linear, *Bottom*: branched). The best candidate for each UA model is highlighted in green.

	UA Models				
Candidates	**Dimethyl ether**	**Dimethyl sulfide**	**Acetonitrile**	**Propene**	**Propane**
Dimethyl ether	**0.115**	**38.0**	**129**	**2.77**	**2.65**
Dimethyl sulfide	**47.2**	**0.138**	**313**	**77.7**	**21.7**
Acetonitrile	**52.7**	**120**	**0.530**	**46.9**	**60.1**
Propene	**4.08**	**42.3**	**85.7**	**0.114**	**3.92**
Propane	**3.84**	**20.9**	**154**	**10.5**	**0.253**
Dimethylamine	**0.921**	**30.9**	**143**	**4.34**	**1.70**
Perfluorohexane	**611**	**223**	**1270**	**802**	**427**
Hexane	**66.1**	**38.1**	**482**	**63.0**	**63.1**
Ethylamine	**2.80**	**27.0**	**144**	**5.05**	**3.01**
Dichloromethylene	**40.7**	**4.09**	**254**	**64.7**	**16.8**

	UA Models			
Candidates	**Acetone**	**Isobutane**	**Isobutylene**	**DMSO**
Acetone	**1999**	**14276**	**5244**	**63994**
Isobutane	**34853**	**1906**	**15919**	**29826**
Isobutylene	**7348**	**68089**	**135**	**50300**
DMSO	**55151**	**22928**	**48257**	**204**

It is apparent from Table 7.1: Table of Φ values for candidate CS-molecules fitting onto a set of UA models of two different topologies (*Top*: linear, *Bottom*: branched). The best candidate for each UA model is highlighted in green.

	UA Models				
Candidates	**Dimethyl ether**	**Dimethyl sulfide**	**Acetonitrile**	**Propene**	**Propane**
Dimethyl ether	**0.115**	**38.0**	**129**	**2.77**	**2.65**
Dimethyl sulfide	**47.2**	**0.138**	**313**	**77.7**	**21.7**

Acetonitrile	52.7	120	0.530	46.9	60.1
Propene	4.08	42.3	85.7	0.114	3.92
Propane	3.84	20.9	154	10.5	0.253
Dimethylamine	0.921	30.9	143	4.34	1.70
Perfluorohexane	611	223	1270	802	427
Hexane	66.1	38.1	482	63.0	63.1
Ethylamine	2.80	27.0	144	5.05	3.01
Dichloromethylene	40.7	4.09	254	64.7	16.8

	UA Models			
Candidates	Acetone	Isobutane	Isobutylene	DMSO
Acetone	1999	14276	5244	63994
Isobutane	34853	1906	15919	29826
Isobutylene	7348	68089	135	50300
DMSO	55151	22928	48257	204

that similar CS-molecules show similar values for Φ. Taking the candidate molecule dimethyl ether as an example, it maps best onto its respective UA model, however one observes that it maps equally well onto the UA models for propane and propene, as these molecules are very similar to each other. In comparison, dimethyl sulfide only maps well onto one molecule, the UA model of dimethyl sulfide. This is due to the long bond length between the sulfur and carbon in the molecule (1.81Å) which precludes dimethyl sulfide from mapping well onto any other UA model. The only other molecule that comes close to map well the UA model of dimethyl sulfide is dichloromethylene, which has long bonds between carbon and chlorine (1.781Å), similar to the carbon-sulfur bond in dimethyl sulfide.

7.3.2 Polymer CG models

To further validate our methodology, we perform a similar comparison as in the previous section (7.3.1) but using now CG models of polymers as testbeds. While the difference

between the AA and UA force-field representation of molecules was minimal, typical CG polymer models average out more atomistic details as one or multiple monomers are mapped onto single beads, hence making the process of finding the optimal AA molecule that fits onto the CG molecule more challenging. Due to the larger DOC, we use the second form of Φ [Eq. 5.7] which assesses differences in g(r). As before, we propose a group of candidate molecules (simulated using the UA model) to fit onto the CG models and calculate Φ for each of them. The two CG models used are the popular MARTINI CG model for PEO, and a CG model for PS which places the center of mass of each CG bead onto the backbone carbon connected to the phenyl ring. This model for PS uses tabular potentials, so any mapping which creates long bonds between CG bead will cause the algorithm to crash. To circumvent this issue, we bypassed the automated algorithm and manually created the mappings for the CG PS model, only using polymer backbone atoms in the CG group. This is done to focus on the ability of our ICG methodology to pick the correct candidates.

As shown in Table 7.2: Table of Φ values for Candidate polymer CS-molecules (5-mers) fitting onto 2 different previously developed CG models. The best candidates for each CG model are highlighted in green.

	CG Polymer Models	
Candidates	MARTINI PEO	CG PS
PEO	0.41	N/A
PS	2.03	0.46
PI	0.91	N/A
PMMA	2.30	1.38
PP	1.16	1.92
PE	0.28	6.87

, the best candidate for the CG model of PS is expectedly found to be the AA model of PS. For the MARTINI CG model of PEO, however, both PE and PEO are found to be the best candidates, with a slightly worse fit for PEO (higher Φ). This inconsistency can be attributed to the fact that the MARTINI PEO CG model was constructed and parameterized for the simulation of biological systems in an aqueous environment, while our simulations describe the melt behavior of PEO. This issue notwithstanding, PEO and PE are comparatively the best candidates for this model, a reflection of the fact that the GC-to-AA mapping problem need not have a unique solution. Similar to what was observed in Table 7.1: Table of Φ values for candidate CS-molecules fitting onto a set of UA models of two different topologies (*Top*: linear, *Bottom*: branched). The best candidate for each UA model is highlighted in green.

	UA Models				
Candidates	Dimethyl ether	Dimethyl sulfide	Acetonitrile	Propene	Propane
Dimethyl ether	0.115	38.0	129	2.77	2.65
Dimethyl sulfide	47.2	0.138	313	77.7	21.7
Acetonitrile	52.7	120	0.530	46.9	60.1
Propene	4.08	42.3	85.7	0.114	3.92
Propane	3.84	20.9	154	10.5	0.253
Dimethylamine	0.921	30.9	143	4.34	1.70
Perfluorohexane	611	223	1270	802	427
Hexane	66.1	38.1	482	63.0	63.1
Ethylamine	2.80	27.0	144	5.05	3.01
Dichloromethylene	40.7	4.09	254	64.7	16.8

	UA Models			
Candidates	Acetone	Isobutane	Isobutylene	DMSO
Acetone	1999	14276	5244	63994
Isobutane	34853	1906	15919	29826
Isobutylene	7348	68089	135	50300

DMSO	55151	22928	48257	204

, candidate molecules of similar chemical structure have comparable viability to fit a given CG model. For the CG PS model in particular, molecules with a monomer that has two carbon atoms in the backbone and a side chain all show similar Φ scores, while molecules such as PEO and PI fail to produce a solution due to the high energetic penalties in the CG model associated with the mismatch of bond lengths.

Table 7.2: Table of Φ values for Candidate polymer CS-molecules (5-mers) fitting onto 2 different previously developed CG models. The best candidates for each CG model are highlighted in green.

	CG Polymer Models	
Candidates	MARTINI PEO	CG PS
PEO	0.41	N/A
PS	2.03	0.46
PI	0.91	N/A
PMMA	2.30	1.38
PP	1.16	1.92
PE	0.28	6.87

7.3.3 Generic CG models

Having confirmed the validity of our methodology when looking at CG models having unscaled length units, we now examine the case of generic CG models originally developed using generically scaled length units. For this purpose, we test our methodology for the Kremer-Grest (KG) bead-spring model of a linear 10-mer AB diblock copolymer (DBC) (5 A beads, 5 B beads). This model has been shown to correctly represent the expected trends in equilibrium and transport properties of polymeric systems[138,186] and several KG-based variants have been used in many simulation studies[149,151,152]. Three cases of the model

are examined: (i) the base case of a homonuclear, flexible DBC ("flexible" model), (ii) a similar DBC but with an added bending potential to increase the persistence length ("stiff" model), and (iii) the base DBC except the B beads are larger in diameter ($\sigma \rightarrow 1.5\sigma$) ("size-asymmetric" model). The forcefields for these systems are detailed in the SI (section 7.6.3). All three cases reflect common modifications to the KG model and represent distinct challenges to our methodology. We used K=1 in Eq. 7.5 when finding the optimal mapping for each candidate to have penalties on the same thermal energy scale as the LJ interactions.

The candidate CS-molecules (chemistries) are chosen from typical monomers used in DBCs and are constructed so that one chemical monomer maps onto a single bead in the KG model. Figure 7.4 is a graph bar showing the fitting scores for each candidate against the three different models in question.

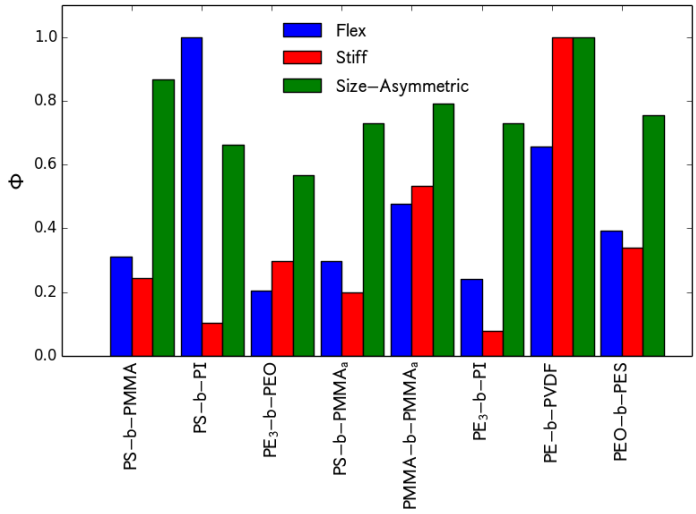

Figure 7.4: Φ values for candidate DBC CS-molecules (10-mers, 5 monomers of each block) fitting onto 3 different forms of the KG model. $PMMA_a$ is the same as normal PMMA, except that the oxygens are replaced with carbons. PE_3 is the same as normal PE, except that a monomer is considered to have 3 CH_2 groups instead of 2.

For the flexible model, PE_3-b-PEO is found to be the best candidate, which is appropriate given that both of the constituent polymers have the lowest persistence lengths among those tested. For the stiff model, PE_3-b-PI ranks as the best candidate, likely due to the increased backbone stiffness of the PI block which is attained without explicit bulky side-groups (e.g., PS, PMMA, $PMMA_a$) to keep the relative size between PE_3 beads and PI beads constant. PE_3-b-PEO ranks as a relatively worse candidate for the stiff model due to the flexibility of its blocks. For the size-asymmetric model, many of the proposed candidates show similar values of Φ with PE_3-b-PEO again having the lowest (best) score. This aligns with idea that most DBCs have blocks where the monomers are of differing volumes and so they should all fit a model with moderate size-asymmetry equally well. In this case, PE-b-PEO ranks as the best candidate likely due to the flexibility of the blocks, which also made it the best candidate for the flexible model.

As alluded to in previous sections where the PE monomer was treated as either 2 or 3 CH_2 groups, a particular issue with the mapping of polymeric systems onto unscaled models is not knowing how many monomers to map to a single bead in the CG model. Models with non-penetrable beads like the KG model tend to be better represented with fewer monomers per bead, while models with soft non-bonded interactions such as the one typically used with the Dissipative Particle Dynamics (DPD) model, tend to be associated with more monomers per bead. To verify this trend, we simulated CG homonuclear 5-mers using these two different models and tried to map polymers of varying DOC onto them, where DOC is now quantified as the number of monomers in the CS candidate molecule being mapped onto 1 CG bead (see SI section 7.6.3.3 for DPD model). For the KG model, it is found that the

minimum in Φ for PEO and PI occurs for DOC = 1 and monotonically increases with DOC (see Figure 7.5, dashed lines), while for PE the Φ minimum occurs for DOC=2. This suggests that for PE the most suitable scale of KG coarse graining would correspond to ~4 backbone atoms. Indeed, the minimum of Φ for each molecule indicates that the best degree of coarse-graining corresponds to 3, 4, and 4 backbone atoms for PEO, PI, and PE, respectively. The persistence length of the flexible KG model is ~1.5σ[202], which suggests that the contour length of 4 backbone atoms in PE should be similar to the persistence length of PE[166]. The contour length of 4 C-C bonds is approximately 6Å, which is close to the experimental value of 6.5Å. Similar agreement is found when considering the optimal DOC length scale we find and the persistence lengths for PEO[105] and PI[57]. This consistency will translate into a suitable mapping of several structural properties of the CS polymers known to correlate with the persistence length, such as the scaling with molecular weight of the average end-to-end distance and radius of gyration. Our results agree with previous studies suggesting the KG model is a suitable CG representation of many polymers.

In contrast, for the DPD model (Figure 7.5, solid lines) the Φ curves show a minimum at around DOC=3-4. These larger values are consistent with the softer potentials used in the DPD model. While a DPD model bead was originally intended to be able to represent a broad spectrum of DOC, studies[20,231] have shown that there are limitations.

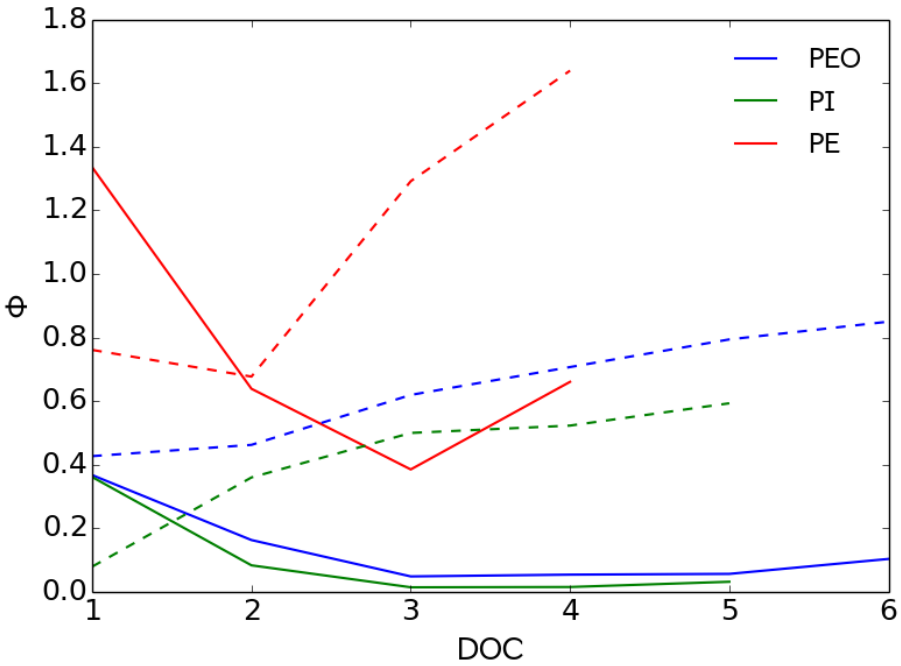

Figure 7.5: Plot of Φ as a function of DOC for different polymers. Solid and dotted lines correspond to the DPD and KG models, respectively.

7.4 Conclusions

We have proposed algorithms to tackle an outstanding problem of molecular modeling concerning the inverse coarse-graining (ICG) process to find the best (chemically specific) molecules to map onto a known CG model of interest. To achieve this, several tools have been developed to facilitate the necessary steps in automating the ICG process, including the determination of the optimal CG mapping of a candidate molecule onto the known CG model, and a correspondence algorithm to uniquely determine if a given CG mapping results in a CG molecule whose topology is consistent with that of the desired CG molecule. While most available CG tools attempt to determine an "optimal" CG mapping in the absence of constrains on the topology of the CG molecule, both of the tools proposed here address the previously unresolved problem of identifying specific chemistries that best satisfy a pre-established CG molecule.

Our new mapping tools are used in conjunction with objective functions to screen for the best CG model fit from a pool of candidate molecules. Our objective functions are constructed based on metrics embodying energetic and structural metrics (including one previously used in the Iterative Boltzmann Inversion process). The methodology shows significant sensitivity of the optimal candidate identified to changes made to a CG model (Figure 7.4). Additionally, the method is sensitive to the degree of coarse-graining adopted, a property that was leveraged to quantify the optimal length scale of coarse-graining for two common generic CG model classes (KG bead-spring and DPD models).

Looking forward, we envision three main ways by which the performance of the proposed ICG methodology can be improved. First, the correspondence algorithm can be leveraged to fit predefined bead structures onto candidate molecules. These predefined structures would come from a database of literature-extracted data, or from a library of structural motifs (fragment library) collected from other users employing this method on their own molecules. This would allow the generation of initial structures that are closer to the optimal one by taking advantage of atom groupings and mappings already known to work well. Second, more advanced objective functions can be developed to help better discriminate between candidates, and to reduce the number of candidates that are kept for the next generation. Third, implementing a fully automated process will expedite the ICG process to allow it to continually advance without requiring user input. Studies are already underway in some of these fronts. The proposed methodology is relatively simple and robust and while further refinements are likely needed, it offers a suitable platform to tackle the challenging problem of Inverse Coarse-Graining.

7.5 Acknowledgments

This work was supported by the National Science Foundation Awards CMMI 1435852 and DMREF Award Number 1629369. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1053575.

7.6 Supplementary Information

7.6.1 More Complex Example of Found Schemes

In Figure 7.6 we provide an example of "optimal mapping" for a complex molecule. We previously studied the molecule, poly(3-methoxyethoxyethoxymethyl), P3MEEMT[46] via both AA and CG models. To validate our proposed algorithms, we simulated a single chain in vacuum and used the objective function (Eq. 7.5 in the main text with K= 0) to modify our mappings with the target CG model being that used in our previous work. With the compression algorithm (detailed below) the monomer structure was identified so the problem of finding the optimal mapping became numerically tractable. Comparing the "optimal" scheme thus found to the "correct" scheme, the scheme which is assumed to be correct based on CG model, there is only a small discrepancy which could be attributed to the fact that the CG model used to represent P3MEEMT was developed for a different, albeit similar molecule, poly(3-hexylthiophene) (P3HT). The equilibrium length of the carbon-carbon bond (1.54Å) is longer than the carbon-oxygen bond (1.41Å) and the equilibrium bond angle for the carbon-carbon-carbon angle (112.7°) is larger than the bond angle for the carbon-oxygen-carbon angle (109.5°). Both these contribute to the side chain of P3MEEMT having a shorter end-to-end length than for a fully alkyl side chain as in P3HT.

The “optimal” mapping increases the average distance between the center of mass of the side chain groups, to closer mimic that of an alkyl side chain, to better match the CG model.

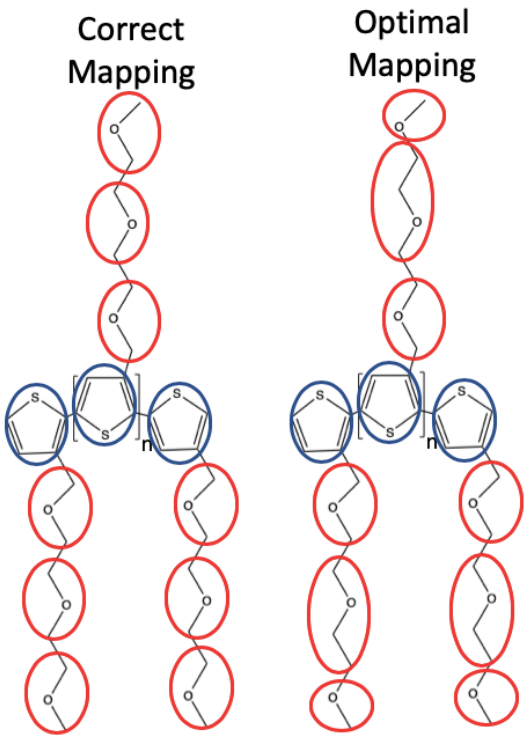

Figure 7.6: Comparison between the “optimal” mapping found by our algorithm, to that of the “correct” mapping of the CG model used in our previous work.[1] The group colorings is to distinguish between the backbone and side chains of P3MEEMT.

7.6.2 Code for compression

With each “smile” generated by our algorithm (see section 7.2.3), a string compression is applied, similar to the methodology used in the zip file format [236,237] as shown in Figure 7.7. The main factor controlling the compression is the “stride”, i.e., the size of the smile fragments that are being compared. Starting from the first character in the smile string, the fragment of the given stride size called the compression pattern, CP, is compared to the adjacent fragment of the same stride size in the smile, and if these two fragments are the same, then a compression is possible and the comparison of adjacent fragments continues until a fragment doesn’t match the repeating pattern. All these repeating patterns are

compressed into a new character, which drastically reduces the total information to process. If no patterns are found with a given fragment, the position along the smile is shifted by one, and a new CP is created. This process is repeated until the end of the string is reached.

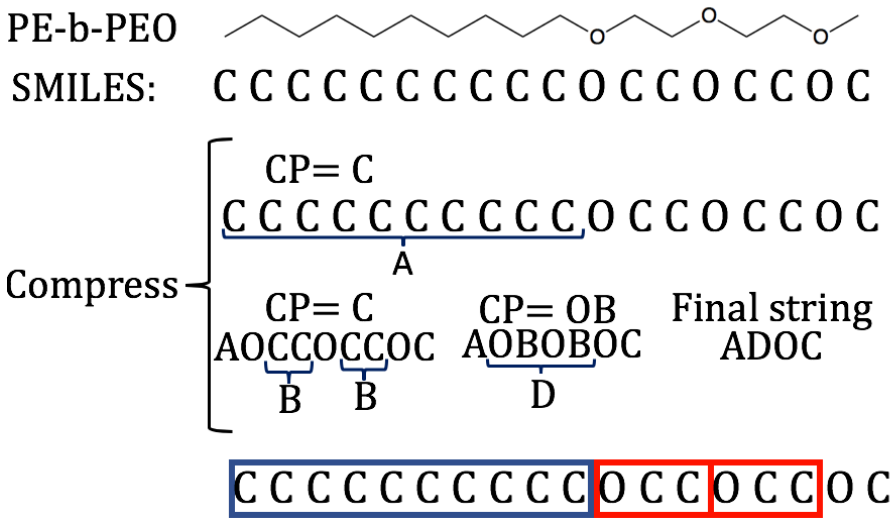

Figure 7.7: A schematic description of the compression strategy. The molecule is turned into the SMILES representation, and undergoes a compression strategy where a compression pattern (CP) is compared to the smile string to see if there are adjacent repeating instances of the CP. If so, the string representing the string of the repeating CP is assigned a new alias and is replaced in the smile string. Once no further compression occurs, the compression is inverted and the overall patterns are found.

The initial compression is done with a stride size of 1, i.e., adjacent atoms of the same type in the string are compressed. Following this, a loop is initiated which "rasters" a copy of the compressed string over itself, illustrated in Figure 7.8. For a given amount the copy is shifted compared to the original smile (Raster amount) N_{match} is the number of elements which match. The copy is shifted until the first non-trivial maxima is found in N_{match}. This gives the stride size to be used in the next round of compression, which returns a newly compressed smile. This process is repeated until no further compression is achieved. The compression is then inverted, where atoms contained in a compression pattern involving more than 2 atoms are assigned to a family of "repeating units" (red boxes in Figure 7.7). To respect this

structure, groups involving atoms of a repeat unit should only contain atoms from that single repeat unit, and the atoms of the other repeating units be grouped in the same way. To achieve this, the algorithm for finding the initial schemes as applied to entire molecules is instead applied exclusively to the atoms inside one of the repeating units. The groups returned by this mapping are propagated across all the repeat units. After the repeating units have been mapped, any other atoms involved in a compression are mapped. Subsequently, the remainder of the molecule is mapped.

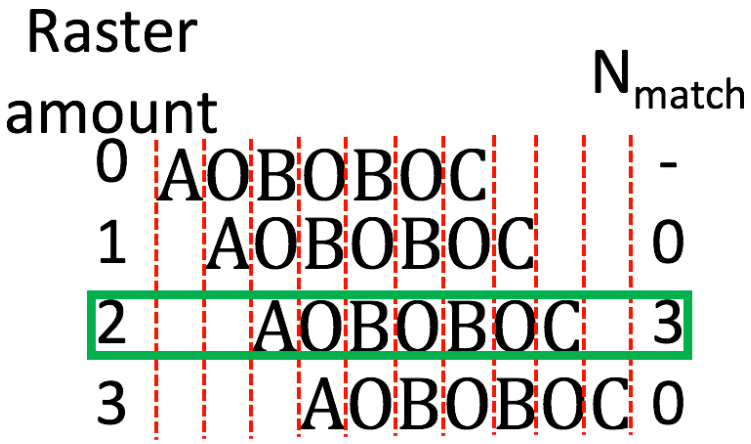

Figure 7.8: Illustration of the rastering approach where a copy of the original string (raster amount=0) is shifted (raster amount>0) and the elements are compared, where N_{match} is the number of matching elements between the original and shifted strings. The first non-trivial maxima in N_{match} is the stride size to be used in the compression algorithm.

When modifying the mapping with repeating units, any modifications made to the groups within one repeating unit should be propagated to all the units. If a merger is proposed between a group inside a repeat unit and a group outside a repeat unit, then the repeat unit that is part of this merger is no longer considered part of the repeating units. Another way for changing the scheme while respecting the repeating units is through “reptation” (Figure 7.9), which involves shifting the borders of the repeating units along the backbone found

using the smile generation algorithm. This is done to allow for a correction in the placement of the repeating units if needed.

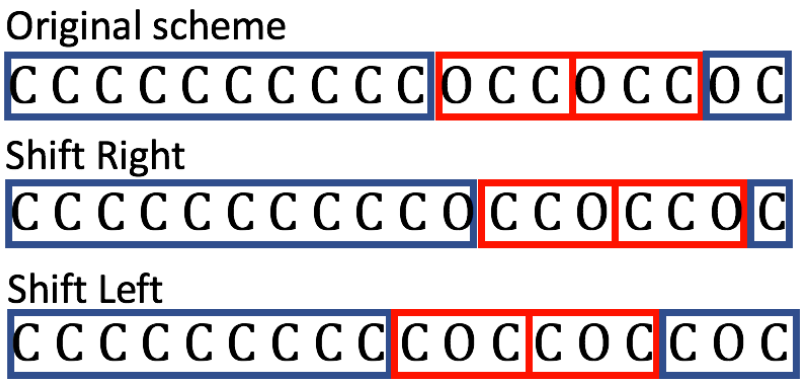

Figure 7.9: Illustration of the reputation method for modifying (please change "schemes" to a more descriptive word). The repeating units can be shifted in one direction or another.

7.6.3 Simulation Methods

7.6.3.1 AA and UA models.

We use the OPLS-AA/UA forcefield.[99] The non-bonded interactions are defined by a combination of Lennard-Jones (LJ) and Coulombic potential energy functions truncated at a cut-off of 10Å. The parameters for homonuclear non-bonded interactions, bond, angle, and dihedral potentials are all taken from [99]. We adopted the OPLS-AA forcefield prescription for the mixing rules for heteronuclear non-bonded interactions and scaling of non-bonded interaction for bonded atoms. The Nosé-Hoover thermostat/barostat was used for Thermostating/Barostating. Time integration of the equations of motions is handled with the velocity-Verlet algorithm with a time step of 1fs.

To prepare these systems, the temperature was initially raised to 600K and the system allowed to equilibrate. We consider a system equilibrated when steady state values are reached for the bending, non-bonded, and bonded potentials, as well as the pressure tensor

components. The system is then cooled at a rate of 15K/ns to 300K and allowed to equilibrate. Once equilibrated, the system is run for approximately 10ns to collect data for analysis.

7.6.3.2 KG Model Variants

The three CG models tested and reported in section 7.3.3 of the main text (Flexible, Stiff, Size-Asymmetric) are variants of the well-known KG model for polymers [72]. To work in reduced LJ units, properties are scaled as follows: $T^*=Tk_b/\varepsilon$, $\tau=t(\varepsilon/(m\sigma_{LJ}^2))^{1/2}$, and ρ=(# of particles)/ σ_{LJ}^3, where T is temperature and ρ is number density, and all scaling variables (σ_{LJ}, m, ε) are set to 1. Bonding was modeled by the standard FENE potential:

$$U_{FENE} = -\frac{K^{ij}\left(R_0^{ij}\right)^2}{2}\ln\left(1-\left(\frac{r}{R_0^{ij}}\right)^2\right)+4\varepsilon\left(\left(\frac{\sigma_{ij}\sigma_{LJ}}{r}\right)^{12}-\left(\frac{\sigma_{ij}\sigma_{LJ}}{r}\right)^{6}+\frac{1}{4}\right) \quad (7.8)$$

where the LJ-like potential term is cut off at $r=2^{1/6}\sigma_{ij}\sigma_{LJ}$ to mimic excluded volume interactions and K^{ij}, R_0^{ij}, and σ_{ij} are the spring force constant and maximum bond extension length, and the interaction scaling distance between a bonded pair i and j, respectively. Table 7.3 gives the parameters for all bond types across all 3 model variants

Table 7.3: Parameters for U_{FENE} for all three KG model variants

Parameter	Flexible	Stiff	Size-Asymmetric
K^{AA} $(\varepsilon/\sigma_{LJ}^2)$	30	30	30
K^{BB} $(\varepsilon/\sigma_{LJ}^2)$	30	30	30
$K^{AB}(\varepsilon/\sigma_{LJ}^2)$	30	30	30

R_0^{AA} (σ_{LJ})	1.5	1.5	1.5
R_0^{BB}(σ_{LJ})	1.5	1.5	2.25
R_0^{AB}(σ_{LJ})	1.5	1.5	1.5
σ_{AA}	1	1	1
σ_{BB}	1	1	1.5
σ_{AB}	1	1	1

For non-bonded interactions, a cut and shifted LJ potential was used with the form:

$$U_{LJ} = 4\varepsilon\varepsilon^{ij}\left(\left(\frac{\sigma_{ij}\sigma_{LJ}}{r}\right)^{12} - \left(\frac{\sigma_{ij}\sigma_{LJ}}{r}\right)^{6} - \left(\frac{\sigma_{ij}\sigma_{LJ}}{r_c^{ij}}\right)^{12} + \left(\frac{\sigma_{ij}\sigma_{LJ}}{r_c^{ij}}\right)^{6}\right) \qquad (7.9)$$

for $r/\sigma_{LJ} \leq r_c^{ij}$ (U_{LJ} = 0 otherwise), where the r_c^{ij} are chosen such that A attract A, B attract B, and A repels B. Table 7.4 gives the parameters for all interaction types across all 3 model variants.

Table 7.4: Parameters for U_{LJ} for all three KG model variants

Parameter	Flexible	Stiff	Size-Asymmetric
ε^{AA}	1	1	1
ε^{BB}	1	1	1
ε^{AB}	1	1	1
σ_{AA}	1	1	1
σ_{BB}	1	1	1.5
σ_{AB}	1	1	1.25
r_c^{AA}(σ_{LJ})	2.5	2.5	2.5

$r_c^{BB}(\sigma_{LJ})$	$2.5\sigma_{LJ}$	$2.5\sigma_{LJ}$	2.5
$r_c^{AB}(\sigma_{LJ})$	$2^{1/6}$	$2^{1/6}$	1.4

For the Stiff model the original KG model is modified by introducing a bending potential that is applied to all beads, which has the form:

$$U_{angle} = k_b T K_{bend}\left(1+\cos(\psi)\right) \qquad (7.10)$$

where K_{bend} is a stiffness constant and ψ is the angle between two consecutive bond vectors. K_{bend} is 4 for the Stiff model and 0 otherwise.

We use 100 10-mer DBC chains (5 A beads, 5 B beads) for every model. For all our simulations we use a constant value of T^*= 1.0 using the Nosé-Hoover thermostat. To prepare the systems, the chains are compressed to the target ρ (0.85 for Flexible and stiff, 0.389 for size-asymmetric), and allowed to equilibrate at constant ρ. Afterwards, the system is run for an additional 10^8 timesteps to collect data. Time integration of the equations of motion is performed using the velocity-Verlet algorithm with timestep τ=.005.

7.6.3.3 DPD Model

The DPD model used in this study is similar to the originally proposed model [74] where the force on a given atom i due to its j neighbors is modeled by:

$$f_i = \sum_{j\neq i} F_{ij}^C + F_{ij}^D + F_{ij}^R + F_{ij}^B \qquad (7.11)$$

where F_{ij}^C, F_{ij}^D, F_{ij}^R, and F_{ij}^B are the conservative, drag, random, and bonded forces, respectively. The conservative force is modeled as:

$$F_{ij}^{C} = \begin{cases} 25(1 - r_{ij})\hat{\boldsymbol{r}}_{ij} & (r_{ij} < 1) \\ 0 & (r_{ij} \geq 1) \end{cases} \quad (7.12)$$

$\boldsymbol{r}_{ij} = r_i - r_j, r_{ij} = |\boldsymbol{r}_{ij}|, \hat{\boldsymbol{r}}_{ij} = {\boldsymbol{r}_{ij}}/{r_{ij}}$. For any two particles bonded to each other, the bonded force is modeled as:

$$F_{ij}^{B} = -2\hat{\boldsymbol{r}}_{ij} \quad (7.13)$$

otherwise $F_{ij}^{B} = 0$. The potentials of F_{ij}^{D}, and F_{ij}^{R} and their parameters are identical to those used in the original model. A Langevin thermostat was used to keep T^{*}=1. Similar to the equilibration methodology for the KG variants, 200 5-mers were compressed to ρ= 3 and allowed to equilibrate. Thereafter, the system is run for an additional 10^{8} timesteps to collect data. Time integration of the equations of motions was performed using the velocity-Verlet algorithm with timestep τ=.04.

8 Stability of the Gyroid Phase in Rod-Coil Systems via Thermodynamic Integration with Molecular Dynamics

8.1 INTRODUCTION

Block copolymers, BCP, are widely used for their ability to self-assemble into mesophases whose geometrical features make them attractive for various types of applications[193]. Among these mesophases, one of the most elusive and yet most sought after phases is the gyroid, or G phase, a bicontinuous morphology with 3-dimensional periodicity, where the minority phases form two non-interpenetrating networks of tubes and nodes, with each node connected to three tubes[78]. BCPs that form the G phase have many potential applications for use as templates in photovoltaic cells[36,200] and electrochemical devices[2,179], because the uninterrupted 3D connectivity of the G phase block domains is beneficial for superior electronic, and mechanical properties. Additionally, bicontinuous phases have large interfacial area between the majority and minority block, which can be coupled with the block connectivity to exploit interfacial effects for applications in heterogeneous catalysis and adsorption[29,85,137].

The G phase is elusive in typical diblock copolymers (dBCP), as it has a narrow composition-temperature window of stability between the cylinder phase, Cyl, and the lamellar phase[131]. Such elusiveness is compounded in molecular simulations, since morphologies with 3D periodicity like the G and other bicontinuous phases, cannot rotate in the simulation box (like the Cyl and lamellar phases), so that any dimensional mismatch between the box size l_{box} with the periodic unit cell will cause internal strains that preclude

such phase from forming. Hence, the formation of the G phase is very sensitive to the proper choice of l_{box}[129].

Different strategies have been used to stabilize the G phase, such as the addition of a homopolymer of the block that forms the bicontinuous networks to fill in the network nodes and relieve chain stretching inside these nodes, a key reason for the limited stability of the G phase relative to other phases[128]. A similar stabilizing effect can also be realized by the addition of a short symmetric diblock "surfactant"[156]. Small molecule solvents can also be used to enhance the access to the G phase as previously demonstrated via thin film vapor annealing with two solvents, each selective for one of the two blocks[155].

While the G phase is primarily observed in DBC systems involving flexible polymer chains, it has been predicted to occur with molecules in which one block is a rigid rod[205]; for instance, in rod-coil BCPs (or RCBCP) composed of a rigid or semi-flexible rod block with persistence length $\gtrsim$10 monomers and a flexible coil block. If the rod forms the minority phase of the G, it eliminates any possible entropic penalty incurred due to nodal chain stretching; however, semi-flexible blocks tend to also form nematic and smectic phases like those observed in lyotropic and thermotropic liquid crystals. Hence, the rod-block must be short enough that a smectic phase is not favored, but this could lead to long coil blocks filling the rest of the volume and favoring micellar phases. To address this short-rod constraint, a previous study[82] examined the phase behavior of a RCBCP system with a solvent selective for the coil block, allowing for the coil to swell and occupy a larger volume fraction of the system despite its short length. Indeed, that simulation study[82] did find

the G phase in RCBCPs but only for a limited subset of parameter space. In this work, a similar RCBCP model has been studied and the G phase was found to spontaneously appear in a different region of phase space than originally reported. However, when simulating the spontaneous formation of ordered phases, different competing phases can be observed over multiple runs of the same state point. At conditions when such competing phases occur, their free energies (relative to a reference state), ΔF, must be used to unambiguously assign phase stability.

Among the different methods available to evaluate ΔF values, thermodynamic integration, TI, is particularly suitable to study ordered phases (for which other methods like those based on particle insertions show large systematic errors[156]). Indeed, TI is an appealing option here due to its generality, ease of implementation, and typically high accuracy. However, the application of TI requires the use of special integration pathways when a first-order transition is involved, as directly crossing it will cause large hysteresis, rendering the path non-reversible and the calculated ΔF value inaccurate. Previous Monte Carlo simulations of flexible dBCPs dealing with the stability of different bicontinuous morphologies demonstrated an implementation of TI where a guiding field[141] was used to gradually order dBCP chains into a desired phase, replacing the first order transition with a continuous transition, and allowing for accurate calculation of ΔF [157]. The guiding fields used in Ref. [141] are discontinuous, so we seek to adapt this method to molecular dynamics and apply it to RCBCP to outline the region of composition-density space where the G is the stable phase. Additionally, we aim to apply of this methodology to more complex molecules comprising coil and rod blocks, such as a bolaamphiphilic system having a side coil chain.

8.2 SIMULATION MODEL AND METHODS

Details of the RCBCP model (the main target of this study) and TI scheme are presented in the following sub-sections, while details of the Bolaamphiphile model are succinctly presented later, in the corresponding Results Section (8.3).

8.2.1 Rod-Coil Model

The polymers, depicted in Figure 8.1, are modeled as bead-spring chains[71], with bonds modeled by a FENE potential:

$$U_{\mathrm{FENE}} = -K\frac{R_0^2}{2}\ln\left(1-\left(\frac{r}{R_0}\right)^2\right) + 4\varepsilon\left(\left(\frac{\sigma_{\mathrm{LJ}}}{r}\right)^{12} - \left(\frac{\sigma_{\mathrm{LJ}}}{r}\right)^{6} + \frac{1}{4}\right) \qquad (8.1)$$

where excluded volume interactions are described by a Lennard-Jones (LJ) like potential term, that is cut off at $r=r_c=2^{1/6}\sigma_{\mathrm{LJ}}$. The spring constant is $K= 30\varepsilon/\sigma_{\mathrm{LJ}}^2$ and the maximum bond extension is $R_0 = 1.5\sigma_{\mathrm{LJ}}$, which describe a stiff bond whose equilibrium length is approximately the bead diameter ($\sim 0.97\sigma_{\mathrm{LJ}}$). For non-bonded interactions, there are 2 unique bead types: A(B) beads represent the rod(coil) block. The RCBCP chains are composed of a rod block of 5 beads, and a coil block of 4 beads. For non-bonded interactions, a shifted force LJ potential, is used of the form:

$$U_{\mathrm{SF}} = U_{\mathrm{LJ}}(r) - U_{\mathrm{LJ}}(r_c) - s*(r-r_c)\left(\frac{dU_{\mathrm{LJ}}(r)}{dr}\right)_{r=r_c};\ U_{\mathrm{LJ}}(r) = 4\varepsilon^{\mathrm{i,j}}\left(\left(\frac{\sigma_{\mathrm{LJ}}}{r}\right)^{12} - \left(\frac{\sigma_{\mathrm{LJ}}}{r}\right)^{6}\right) \qquad (8.2)$$

for $r \leq r_c$ (U_{SF} = 0 otherwise), where $\varepsilon^{i,j}$=1, s=1, and r_c=2.5σ_{LJ} for i=j=A, and $\varepsilon^{i,j}$=1, s=0, and r_c=$2^{1/6}\sigma_{LJ}$ otherwise. Hereafter, we refer to the interaction energy between type i and j beads as U_{ij}.

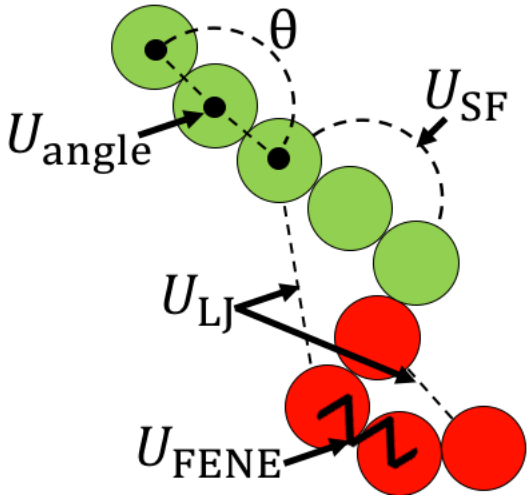

Figure 8.1: A depiction of the RCBCP molecules. A type beads (green) represent the rigid block, where an angle potential induces stiffness. B type beads (red) constitute the flexible block.

Chain stiffness is induced by a bending potential that is applied to successive bonds formed by A type beads only, and has the form:

$$U_{\text{angle}} = k_b T K_{\text{bend}}(1 + \cos\theta) \qquad (8.3)$$

where K_{bend} is a stiffness parameter and θ is the angle between two consecutive bond vectors; unless otherwise noted, K_{bend}=100, which makes the A-block chain approach the rigid rod limit (persistence length≈100σ)[55]. Reported properties are scaled as follows: T^*=Tk_b/ε, τ=$t(\varepsilon/(m\sigma_{LJ}^2))^{1/2}$, and ρ = (# of particles)/ σ_{LJ}^3, where T is temperature, ϱ is number density, and all scaling variables (σ_{LJ}, m, ε) are set to 1. The equations of motion for all simulations are integrated using a velocity Verlet algorithm, with a timestep of 0.005τ and temperature is controlled using a Langevin thermostat.

8.2.2 System preparation and characterization

Simulating the spontaneous formation of ordered phases begins with equilibrating the system for 2×10^6 timesteps with $r_c=2^{1/6}$ for A-A interactions, and $T^*=T^*_{target}+0.5$, where T^*_{target} is the final T^* of interest. Subsequently, r_c for the A-A interaction is set to $2.5\sigma_{LJ}$, and the system is linearly cooled to T^*_{target} over 2×10^7 timesteps. T^* is then held constant while the system equilibrates, a process that is monitored by observing the convergence of $U_{\text{non-bonded}}$, U_{angle}, and the components of the pressure tensor with time. Once equilibrated, the resulting morphology is determined by visual inspection and by examining the structure factor, $S(\mathbf{q})$, defined as:

$$S_{\mathrm{p}}(\boldsymbol{q}) = \frac{\left(\sum_{\mathrm{p}j=\mathrm{p}} \cos \boldsymbol{q}\cdot\boldsymbol{r}_j\right)^2+\left(\sum_{\mathrm{p}j=\mathrm{p}} \sin \boldsymbol{q}\cdot\boldsymbol{r}_j\right)^2}{N_{\mathrm{p}}} \qquad (8.4)$$

where $\boldsymbol{q}$ is the wave vector, $\boldsymbol{r}_j$ is the position of particle j of type p and N_p is the number of type p particles (the rod block).

To characterize the orientational order of the rod-like blocks, we calculate the local distribution of local values the nematic order parameter, P_2, defined as the largest eigenvalue of the ordering tensor matrix:

$$Q_{ij} = \frac{1}{N}\sum_{n=1}^{N}\frac{3}{2}\hat{u}_{in}\hat{u}_{jn} - \frac{\delta_{ij}}{2} \quad (8.5)$$

where $\hat{u}_{in}$ is the unit vector pointing along bond n in direction i, N is the number of bonds considered, and δ_{ij} is the Kroncker delta.

8.2.3 Construction of guiding field

To conduct the TI calculation, a guiding field must be constructed that defines the shape of the desired mesophase, i.e., the spatial distribution of the two blocks. We represent this field by a group of "ghost" particles, denoted as type C, which are held fixed on a cubic grid

molded in the shape of the target mesophase, but interact with type A beads through a Gaussian potential[163]:

$$U_{\text{Gauss}} = -ae^{-br^2} \qquad (8.6)$$

for r<2.5σ_{LJ} (U_{Gauss}=0, otherwise) where a =2, and b =2. The value of a is chosen to be large enough to induce the assembly of the target mesophase but small enough to the number of

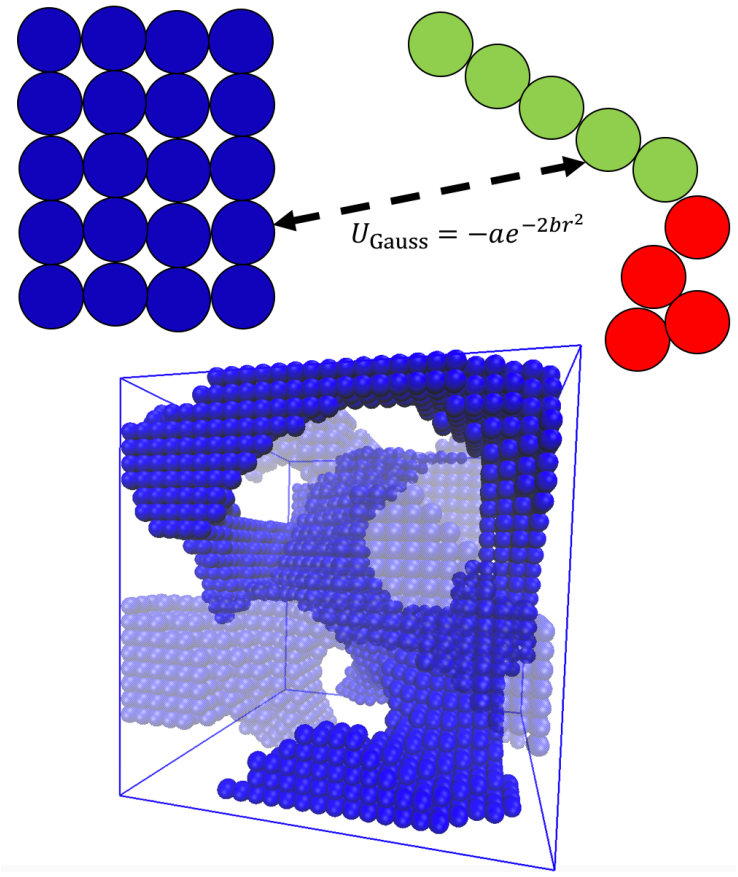

Figure 8.2: A schematic of how RCBCP interact with the field. The C type beads(blue) which make up the field are kept at fixed positions in the shape of the morphology of interest, and only interact with rod segment through a Gaussian potential. The bottom panel depicts a sample guiding field for a G phase where beads represent the Gaussian site locations and one of the two networks is rendered translucent for clarity.

steps needed to turn off the field low. The value of b is chosen such that the mean-field experienced by the rod while inside the field is nearly constant to allow for uninhibited

motion. Figure 8.2 shows a depiction of the field. While the shape of the external field could be described via a continuous function of positions[163] having a special smoothing function across the inter-block interface to allow computation of forces and hence the implementation of standard MD, the use of ghost particles facilitates the description of fields with complex geometrical shapes like that of the G phase.

For every set of parameters, T^*, ρ, l_{box}, and desired morphology, a unique field is required. Construction begins by calculating the structure factor, $S(\mathbf{q})$, for 10 snapshots of the spontaneously formed morphology at the desired parameters. Each $S(\mathbf{q})$ is shifted by a phase factor to best align the morphologies onto one reference snapshot to account for morphological drift. This strategy for constructing the field is described in more detail in Ref. 19. Once the $S(\mathbf{q})$ functions have been shifted, they are inverted to find the density, ρ_p, which is calculated as[87]:

$$\rho_p(\boldsymbol{r}) = 1 + \sum_{k,|\boldsymbol{q}_k|<2} \beta_k \sin(\boldsymbol{q}_k \cdot \boldsymbol{r} + \theta_k); \quad \beta_k = \frac{2}{N_\tau}\sqrt{S_p(\boldsymbol{q})N}\ ; \quad \theta_k = \tan^{-1}\left(\frac{\sum_{\tau_j=p} \cos \boldsymbol{q}\cdot\boldsymbol{r}_j}{\sum_{\tau_j=p} \sin \boldsymbol{q}\cdot\boldsymbol{r}_j}\right) \quad (8.7)$$

where N is the total number of particles. Density is calculated on a cubic grid with spacing equal to (l_{box}/ceiling(l_{box})) $1\sigma_{LJ}$ spacing. Smaller spacings between grid points would approach a continuous function, but are computationally expensive, while larger spacings result in a field potential which is uneven inside the morphology and may unduly perturb the desired morphology. We choose spacing $\leq\sigma_{LJ}$ to avoid the potential spurious effects of large spacings. We average the inverted density across all snapshots, and apply a cutoff such that if $\rho^*(\boldsymbol{r})>1.1$, a type C particle is placed there. The choice of cutoff is determined empirically; we use the lowest value that still allows removal of grid points that bridge microsegregated domains (e.g., points which would connect the two bicontinuous

networks in the G phase). For cases when the G phase is not formed at the target values of T^*, ρ, and l_{box}, a field is "borrowed" from the nearest state where the G was formed, which assumes that there is no large difference between the morphology's shape for proximal points in parameter space. We note that the free energy calculations are largely insensitive to the precise shape of the field morphology as long as it is able to push the system into the appropriate morphological basin.

8.2.4 Thermodynamic Integration Pathway

To calculate the free energy of assembly, Δ*F*, a three-branch process is followed as outlined in Figure 8.3. In the formalism of TI, Δ*F* is written as:

$$\Delta F = \Delta F_1 + \Delta F_2 + \Delta F_3; \;\; \Delta F_i = \int_0^1 d\lambda_i \left\langle \frac{dU}{d\lambda_i} \right\rangle \quad (8.8)$$

Where U is the function describing the potential energy of the system, and λ_i is a coupling parameter that modulates $\underline{U}$. We define λ as the coupling parameter of the field strength, and η_{ii} as the coupling parameter for the non-bonded interaction between beads of type *i*. These parameters affect U through:

$$U_{AC} = \lambda U_{gauss} \quad (8.9)$$

$$U_{AA} = (0.1 + 0.9\eta_{AA})U_{SF} \quad (8.10)$$

Where U_{ij} is the non-bonded interaction between beads of type i and j. In the initial "Reference" state, the external field is off (λ=0), and the A-A interactions are weakened (η_{AA}=0), resulting in the isotropic state. In the first "branch" of the integration, the field is turned on (λ=0→1) at fixed η, giving:

$$\Delta F_1 = \int_0^1 d\lambda \left\langle \frac{U_{AC}}{\lambda} \right\rangle_{\lambda, \eta_{AA}=0} \quad (8.11)$$

In the second branch the field is kept on (λ=1), and the A-A interactions are turned on (η_{AA}=0→1) giving:

$$\Delta F_2 = \int_0^1 d\eta_{AA}(0.9)\left\langle \frac{U_{AA}}{(0.1+0.9\eta_{AA})}\right\rangle_{\eta_{AA},\lambda=1} \qquad (8.12)$$

In the third branch the A-A interactions are kept on (η_{AA}=1), and the field is turned off (λ=1→0) giving:

$$\Delta F_3 = \int_1^0 -d\lambda \left\langle \frac{U_{AC}}{\lambda}\right\rangle_{\lambda,\eta_{AA}=1} \qquad (8.13)$$

Previous work combined we do in branches two and three into a single branch. The reason for splitting into two distinct branches is to allow chains that leave the field volume to re-enter it with a correct orientation. Indeed keeping the field on while the A-A interactions are turned on will induce chains to minimize the energy of both the field, and A-A interactions, promoting proper orientational alignment.

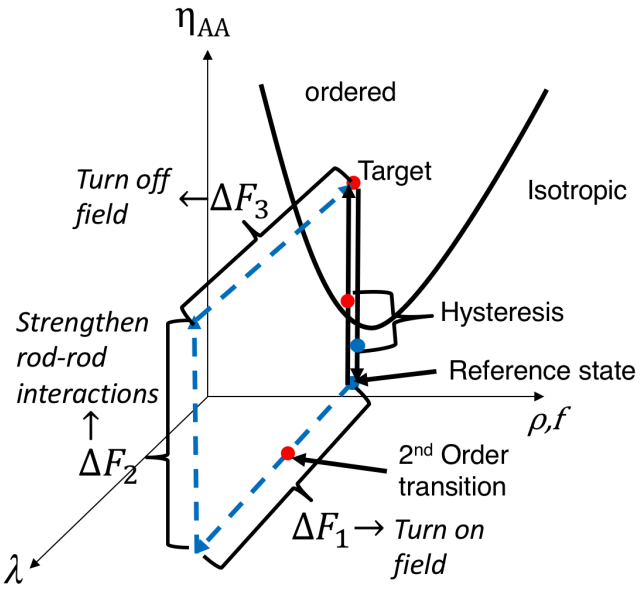

Figure 8.3: Schematics of the branches for the TI pathway. In branch 1 U_{AA}, which represents the rod-rod interactions, is initially weak, so the phase is disordered. The field is turned on, assembling the molecules in the desired phase. In branch 2 the field is kept on, and the rod-rod interactions are restored to full strength. In branch 3 the field is turned off while rod-rod interactions are maintained.

To evaluate the integrals in the free energy equation for each branch, a 10-step Gauss-Legendre quadrature is used[73,178]. For each step in the quadrature, the system is equilibrated for ~2×10^7 timesteps. Only energies calculated after equilibration are considered in the averages used in the TI formulas. Equilibration is defined as having reached time invariance of the integrand in the free energy equation for a given branch. The calculation of ΔF as outlined in Figure 8.3 is for the assembly of the chains into the desired morphology from the reference state at a specific T^*, ρ, and simulation box size, l_{box}. As stated before, the G phase is sensitive to any mismatch between l_{box} and the natural unit cell size which is *a-priori* unknown. As such, ΔF must be calculated for several values of l_{box} to find the minimum ΔF for assembling the given morphology, corresponding to the value of l_{box} closest to the true unit cell size, as we are simulating only single unit cells.

8.3 RESULTS

8.3.1 Linear Rod-Coil Amphiphile Molecules

For our model RCBCP depicted in Figure 8.1, for certain T and ρ, the Cyl, and G phase both spontaneously appear; hence, our ΔF calculations are intended here to discriminate the thermodynamic stability between these two phases. Figure 8.4a shows the values of ΔF for different values of l_{box}, at T^*=1 and ρ= 0.49, showing the dependence of ΔF on l_{box}. For the l_{box} values shown in Figure 8.4a the G and Cyl phases were spontaneously formed; we note that in spontaneously formed Cyl phases the cylinders were always aligned along one of the box diagonals. The minimum ΔF value vs. l_{box} for each phase is henceforth used for comparison of thermodynamic stability at different T* and ρ conditions. Figure 8.4b plots the difference ΔF_G-ΔF_C, showing the expected trend where the stability of G relative to Cyl

increases with increasing ρ. At T^*=1.0 the Cyl becomes the most stable phase for $\rho \leq 0.49$, with the G phase not even forming for $\rho \leq 0.46$, corroborating the prevalence of the Cyl phase. The plotted errorbars are calculated by propagating the error from each step in the TI, which is found using block averages.

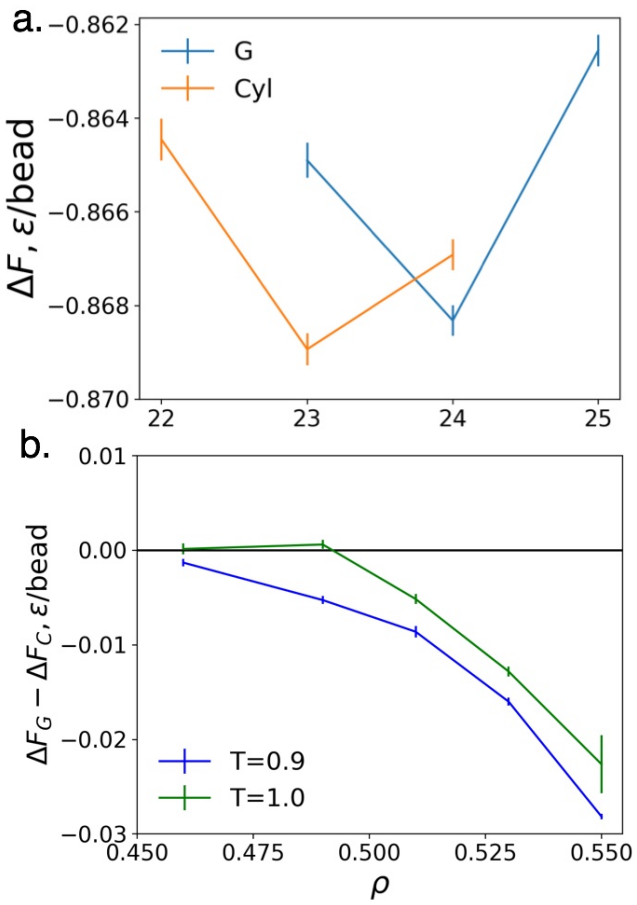

Figure 8.4: a) Variation of ΔF with l_{box} at T^*=1 and ρ= 0.49 showing that ΔF has a minimum at a certain box size which corresponds approximately to the phase's unit cell size. b) Plot of the difference in ΔF between G and Cyl phases. The G phase is increasingly more stable with ρ, while Cyl becomes the more stable phase for T^*=1.0 and ρ < 0.49. The symbol sizes depict the errorbars.

Plotted in Figure 8.5 is a temperature-concentration phase diagram with approximate phase boundaries drawn from a combination of ΔF calculations and observations of spontaneous ordering from numerous simulations. In this phase diagram, an approximate

region of G stability has been established using ΔF calculations. We note that the calculated ΔF values are for the 'driven' morphologies, rather than the spontaneous morphologies. As shown in Figure 8.9, we find that spontaneous G morphologies show a defect which is lacking in the 'driven' morphologies, however, since the defect likely increases the free energy of the phase containing it[183], it follows that the free energy change associated with the 'driven' G phase, ΔF_{driven}, will likely be slightly larger than that of the spontaneously assembled G phase, $\Delta F_{\text{spont.}}$ Unlike the situation for the G phase, no structural difference was detected between Cyl that are formed spontaneously or field-driven, consistent with the absence of topological defects. As such, comparing the ΔF_{Cyl} to ΔF_{driven} will yield accurate discrimination between the two phases.

In previous work for a RCBCP with coil-length of 4 as ours, [82] the authors found that the G phase is observed at $\rho = 0.4011$, $T^* \approx 0.8\text{-}2.0$. We were unable to stabilize the G phase at such low density; instead we observe the G phase for $\rho \approx 0.47\text{-}0.57$, $T^* \approx 0.9$. The simulation algorithms used in the previous and our work have some differences: The previous study used a leapfrog algorithm with a Langevin thermostat for the integration of the equations of motion with a timestep of $0.01\boldsymbol{\tau}$, while we used the Langevin thermostat, and a velocity-Verlet algorithm with a timestep of $0.005\boldsymbol{\tau}$. While equilibrating the system, the previous work employed a box search algorithm which alters the box dimensions but still keeps volume constant[190] while we use a simulation box that has constant dimensions. In our simulations the boxes were cubic but the pressure tensors for G forming systems were found to be isotropic. Perhaps more importantly, the force-field details of the RCBCP models used are not identical: the previous work[82] used rigid body-dynamics for the rods, enforced by the equations of motion for the rotation of linear bodies, while our work

used a stiff bending potential to confer rigidity to the rods. Semi-flexible polymers do not have an infinite persistence length, so our choice of a bending potential to confer rigidity is more realistic. All these differences could be contributing factors to the observed discrepancy on the temperature-density stability region of the G phase.

8.3.2 Bolaamphilic Molecules

We now apply the developed free energy calculation methodology to a more complex type of BCP composed of 3 blocks illustrated in Figure 8.6(a): one rigid rod (B beads in blue), two sticky hydrophilic end groups (A beads in red), and a side chain coil (C beads in purple), which are known as T-shape bolaamphiphiles (henceforth abbreviated as "Bola"). A coarse-grained model for these molecules was previously developed[1] for linear side chains to replicate experimental observations (Figure 8.6b), and has been extended to include a branched side chain denoted "swallow-tail" (ST) (see Figure 8.6c) to access

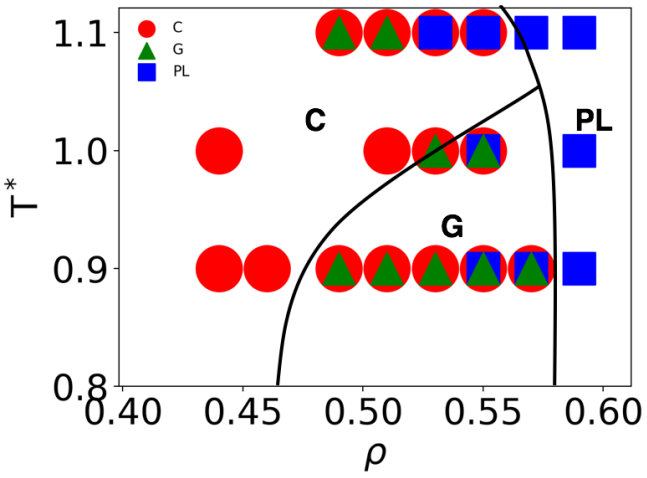

Figure 8.5: Approximate temperature-concentration phase diagram for the rod-coil system studied. Multiple markers on a given point indicate that the corresponding phases were observed in the simulations. The black lines are approximate phase boundaries based on ΔF calculations for the observable phases.

different phases including the G phase[88]. These two studies complement experimental work[31] to show that two molecules with the same number of monomers in the side chain, and only differing in whether the side chain is linear or ST, assemble into different morphologies: either perforated lamellae (PL) or G phase, respectively. Since these molecules have similar χN, the ST side chain must be influencing the entropy of assembly to favor the formation of the G phase. This hypothesis can be tested by comparing ΔF associated with the assembly of the G and PL phases for both the linear and ST side-chain Bolas.

For these simulations, the non-bonded interactions are given by the same U_{SF} function defined in Eq. 8.2 as in the study of RCBCP but with s=0, $\varepsilon^{AA}= \varepsilon^{AB}= \varepsilon^{AC}= \varepsilon^{BC}=\varepsilon^{BB}=1$, and $\varepsilon^{CC}=0.5$, $r_c^{AA}= r_c^{CC}=2\sigma_{LJ}$, and $r_c=2^{1/6}\sigma_{LJ}$ otherwise. The A and B blocks are bonded at a $0.7\sigma_{LJ}$ distance and simulated as a linear rigid body, while the side chain is fully flexible ($K_{bend}=0$) and is bonded using the same FENE potential described in Eq. 8.1. The field potential is modeled using the same Gaussian potential of Eq. 8.6 but with a=1 and the field now interacts with both the A and B blocks. The method for constructing the field is the same as for RCBCP, where instead of the rod, the structure factors are computed with respect to the entire rigid segment (i.e., A+B blocks).

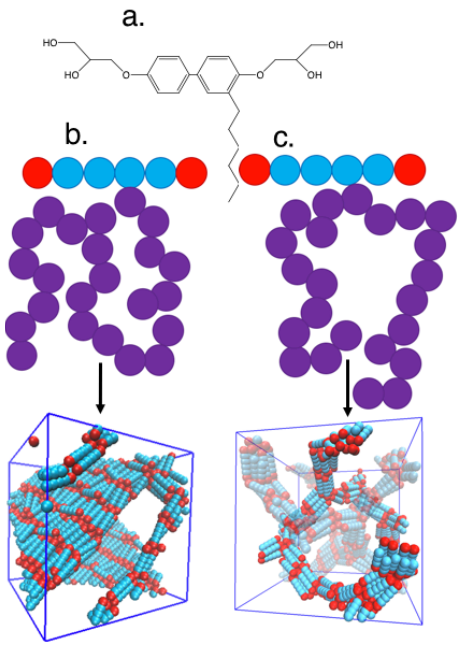

Figure 8.6: a) Sample bolaamphiphile molecule with linear side chain. b/c) Coarse grained model for a molecule with a linear/swallow-tail side chain that forms perforated lamellae/G phase. In the coarse-grained models the A, B, and C type beads are colored in red, light blue, and purple. In the phase snapshots the side chain (C beads) are not shown and one of the two networks of the G phase is rendered translucent for clarity.

The specific molecules of interest have side chains with a total of 29 beads. The systems are simulated at T^*=0.5, ρ=1.3087, and l_{box}=21.694σ_{LJ} , with 300 molecules (conditions at which the G phase forms spontaneously[30]) for the ΔF calculations. This choice of l_{box} is appropriate for the G phase as experience shows that this phase spontaneously forms only in a very narrow range of l_{box} near the optimal value. It is also an appropriate choice for the PL phase as ΔF shows little sensitivity to $l_{box\,f}$ for irregular morphologies that can rotate in the box. A three-branch method is used again but is modified to accommodate the need to return two blocks (A and C) to their original model parameter values. In the first branch we

begin with $r_c^{AA}=r_c^{CC}=2^{1/6}\sigma_{LJ}$ to make the chains athermal. As with RCBCP, the field is turned on linearly with respect to λ. The equation for ΔF_1 is now:

$$\Delta F_1 = \int_0^1 d\lambda \left\langle \frac{U_{rigid-field}}{\lambda} \right\rangle_\lambda \qquad (8.14)$$

In the second branch, U_{AA} and U_{CC} are now modeled through the sum of two potentials; an attractive potential, $U_{AA/CC}^{att} = U_{AA/CC}$ ($r_c=2\sigma_{LJ}$), and a repulsive potential $U_{AA/CC}^{rep} = U_{AA/CC}$ ($r_c=2^{1/6}\sigma_{LJ}$). During this branch U_{AA}^{att} is turned on while U_{AA}^{rep} is turned off with increasing η_{AA}, allowing for a switch between the truly athermal and thermal regimes. The TI equation is:

$$\Delta F_2 = \int_0^1 d\eta_{AA} \left(\left\langle \frac{U_{AA}^{att}}{\eta_{AA}} \right\rangle_{\eta_{AA},\eta_{CC}=0,\lambda=1} - \left\langle \frac{U_{AA}^{rep}}{1-\eta_{AA}} \right\rangle_{\eta_{AA},\eta_{CC}=0,\lambda=1} \right) \qquad (8.15)$$

In the final branch, U_{CC}^{att} is turned on and U_{CC}^{rep} is turned off with increasing η_{CC} (like branch 2), while $U_{\text{rigid-field}}$ is turned off with respect to λ. This gives:

$$\Delta F_3 = \int_0^1 d\eta_{CC} \left(\left\langle \frac{U_{CC}^{att}}{\eta_{CC}} \right\rangle_{\eta_{CC},\lambda,\eta_{AA}=1} - \left\langle \frac{U_{CC}^{rep}}{1-\eta_{CC}} \right\rangle_{\eta_{CC},\lambda,\eta_{AA}=1} \right) - \int_1^0 d\lambda \left\langle \frac{U_{rigid-field}}{\lambda} \right\rangle_{\eta_{CC},\lambda,\eta_{AA}=1} \quad (8.16)$$

Figure 8.7 graphically depicts how the potentials evolve over the course of the TI path, and comparing to RCBCP, the calculations for branches 1 (turning the field on) and 2 (making rod-rod interactions thermal) are qualitatively similar while in the third branch, besides turning off the field, the C-C interactions are also made thermal (shown with the red axis). This 'late' turning on of the C-C interactions allows side chains to stay more mobile while η_{AA}=0→1, facilitating rearrangement of the rigid segments.

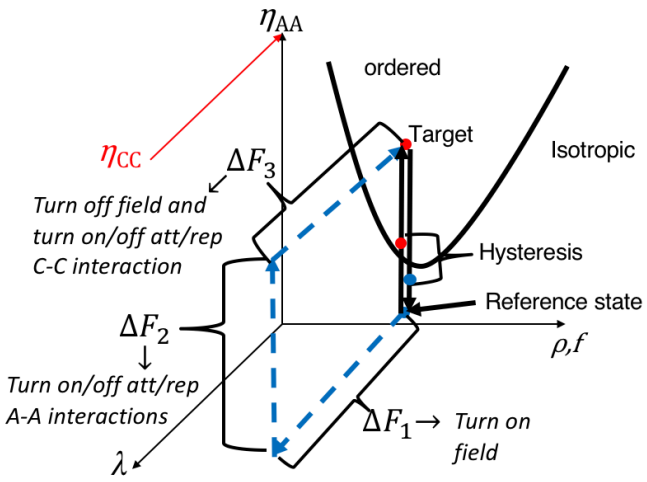

Figure 8.7: Schematic branches of the TI pathway for Bola molecules. In branch 1 all interactions are either repulsive or are turned off, so the system is initially disordered and ends up assembled into the desired phase by turning on the field. In branch 2, the field is kept on and the attractive/repulsive A-A potentials are turned on/off, respectively. In branch 3, A-A interactions are kept and the field is turned off while the attractive/repulsive C-C potentials are turned on/off, respectively (red axis).

Applying our TI methodology to the linear/ST Bola molecules and to the G and PL phases, allowed us to obtain $\Delta\Delta F^{Schain} \equiv \Delta F_G^{Schain} - \Delta F_{PL}^{Schain}$ where Schain denotes side chain type and can be either swallow-tail (ST) or linear (Lin). We obtained $\Delta\Delta F^{\mathrm{ST}}$=-0.0075±0.0015, and $\Delta\Delta F^{\mathrm{Lin}}$=+0.0190±0.0015, indicating that the ST side chain prefers to assemble into the G phase, while the linear side chain prefers the PL phase, in agreement with our hypothesis and prior results[88]. Figure 8.8 plots the cumulative values of $\Delta\Delta F$ as the TI progresses. The linear side chain shows a preference for the PL phase after the first TI branch as captured by $\Delta\Delta F_1^{\mathrm{Schain}} \equiv \Delta F_{G,1}^{Schain} - \Delta F_{PL,1}^{Schain}$; indeed: $\Delta\Delta F_1^{ST}$= 0.00197±0.0005, and $\Delta\Delta F_1^{Lin}$= 0.01902±0.0007. Since in this first branch of the TI the chains are athermal, this difference in free energies is primarily of entropic origin and can be ascribed to the difference in side chain conformational entropy in the target phase.

Previous work on these Bolas[30] argued that the reason why a linear side chain assembles preferentially into the PL phase while the swallow-tail side chain assembles into the G phase is that a flatter interface, such as that in the PL phase, creates a larger steric hindrance to ST chains than to linear chains since the former have larger excluded volume near the "grafting site" (i.e., the minority-block interface of the G/PL). The difference in $\Delta\Delta F_1$ values noted above provides the quantitative origin for the swallow-tail side-chain Bola disfavoring the PL phase in comparison to the linear side-chain Bola.

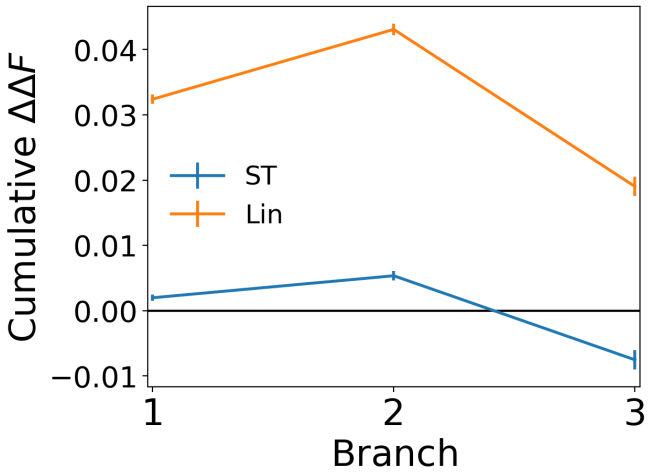

Figure 8.8: Cumulative ΔΔF (G minus PL) through the TI path for the Bola molecules with swallow-tail (ST) and linear (Lin) side-chains. The PL phase has a more negative total ΔF (more stable) than the G phase for the linear side chain (positive ΔΔF), while the G phase has a more negative ΔF than the PL phase for the swallow-tail side chain (negative ΔΔF). The symbol sizes depict the errobars.

8.4 CONCLUSIONS

We have adapted a previously developed MC-based thermodynamic integration methodology for calculating the free energies of ordered phases in polymers to be suitable with MD simulations. The ordering field was constructed as a group of ghost beads, fixed in

space and pixelating the shape of the desired mesophase, that interacted with the real molecules through a tunable Gaussian potential. In constructing the integration path, the sequence in which the field is turned on/off, and the point when the intermolecular interactions are switched from athermal to thermal were selected to overcome some of the difficulties associated with assembling molecules having rigid blocks and more than two types of blocks.

The proposed method was used to discriminate the stability of the Cyl and G phases that spontaneously form in RCBCP at different temperatures and concentrations. Our free energy calculations establish a region of phase stability for the G phase that is inconsistent with findings from a previously reported work[17] on a similar system which didn't involve free energy calculations.

Applying the method to bolaamphiphile systems with differing side-chain architecture (linear or branched) leads to free energy differences whose values show consistency with the findings of previous work[88] indicating that a side chain with a swallow-tail branching favors the G phase over the PL phase, while a linear side chain favors the PL phase.

While the methodology described here is expected to be broadly applicable, specific tuning or extensions could be needed to resolve the challenges posed by some systems. For the Bola system, we adapted the method to account for the need to restore two different types of non-bonded interactions. As the number of interaction types increases, the number of integration path branches required could increase, making the computations more unwieldy. Additionally, for the Bola system the A and B blocks both assembled into the same segregated phase, so the external field didn't have to discriminate between the two. If a desired morphology has more distinctive domains as in the case of ordered phases for

triblock copolymers, then an external field having different types of Ghost particles may have to be designed to drive the different blocks into the target morphology. More generally, the coupling parameter must also be carefully chosen such that the integrand in Eq. 8.8 is easy to evaluate. As an example, the process of changing an intermolecular potential from athermal to thermal can often be implemented in multiple ways. Two ways were explored here: in the first method the strength of interaction was linearly weakened, while in the second method two overlaid potentials representing the thermal and athermal states were switched on/off, respectively. A third way may be to adjust r_{cut} from $2^{1/6}$ (athermal potential) to the desired value; however, it would then be non-trivial to evaluate the expression for the integrand in Eq. 8.7.

Future work will involve applying this TI methodology to other phases in Bola systems so that other questions regarding phase stability and molecular design can be answered more rigorously via free energy calculations. Additionally, the generality of TI and the simplicity of the field implementation can be exploited to study the formation of arbitrarily complex ordered structures with many different types of molecules and multicomponent systems.

8.5 ACKNOWLEDGEMENTS

This work was supported by the National Science Foundation Awards CMMI 1435852 and DMREF Award Number 1629369. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1053575. We thank Prof. Padmanbhan for useful exchanges and Yangyang Sun for providing initial configurations for the bola simulations.

8.6 Supplemental Information

8.6.1 Comparing Spontaneous *vs*. Driven Morphologies

Figure 8.9a (b) shows a snapshot of spontaneously formed Cyl (G) phase (T^*=0.9, ρ=0.53, l_{box}=24σ_{LJ}), and Figure 8.9c shows the distributions of P_2 values for all chains in the configurations from the systems and from a G phase formed by an external-field whose shape was derived from the configuration in Figure 8.9b. The G phase P_2 distributions show greater rod alignment than that for the Cyl phase, reflecting the lamellae-like structures present in the G nodes. The P_2 distributions for the spontaneously formed and field-driven G phases are indistinguishable except around $P_2 \sim 0.45$. This is attributed to a "defect" associated with the mismatch between the local directors of a node and an arm (see example in Figure 8.9b). This means that our field-driven phase does not reproduce the exact structure of the spontaneously formed phase; however, since the defect likely increases the free energy of the phase containing it[88], it follows that the free energy change associated with the 'driven' G phase, ΔF_{driven}, will likely be slightly larger than that of the spontaneously assembled G phase, ΔF_{spont}. In comparing ΔF between the driven G and a competing phase like Cyl, if $\Delta F_{driven} < \Delta F_{spont}$, then there would be no impact on the ability to discern phase stability of G versus Cyl. In the scenario that $\Delta F_{driven} > \Delta F_{spont}$, then the only situation where the phase stability would be uncertain is if $\Delta F_{driven} > \Delta F_{Cyl}$ as then one could not rule out if $\Delta F_{spont} < \Delta F_{Cyl}$. Unlike the situation for the G phase, no structural difference was detected between Cyl that are formed spontaneously or field-driven, consistent with the absence of topological defects. An additional metric to consider is the number of molecules belonging to each G network because an asymmetry should increase the free energy. Taking the difference between the

number of molecules in the two networks and normalizing by the total number of molecules gives average values of 0.023, and 0.005, for the spontaneous, and 'driven' morphologies, respectively. This indicates that molecules are more evenly distributed for 'driven' systems, which should lead to more accurate ΔF values.

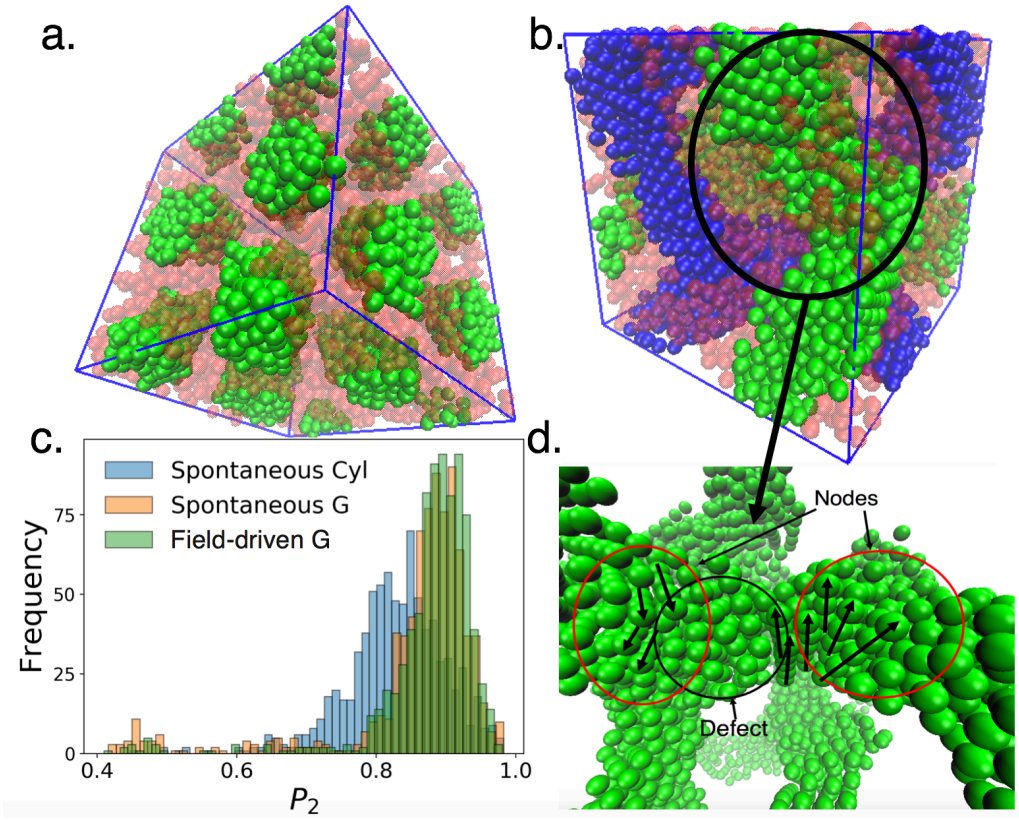

Figure 8.9: Snapshots of spontaneously formed Cyl phase (a) and G phase (b). Coil beads are semitransparent red, while rod beads are green in (a) and green and blue in (b) to discriminate between the two networks. c) Distributions of local P_2 parameter for spontaneously formed Cyl and G phases, and for a field-driven G phase. The two G distributions are largely identical except for region near $P_2 \sim 0.45$ which signal the presence of node-arm defects illustrated in d). Rod misalignments can occur inside an arm as it connects two nodes having distinct rod orientations.

9 Outlook

In this work we aimed to leverage molecular simulation, in particular, Molecular Dynamics (MD), to accelerate the materials discovery process. The clearest example of this is the studies presented in Chapter 2, where we predicted the relative ionic conductivities of two polythiophene derivatives in both the crystalline and amorphous phase before they were synthesized. The experimental results confirm our predictions that just the simple addition of a methylene group in the side chain of a polythiophene derivative can result in a five-fold increase in the ionic conductivity. This study not only is exemplary of the power of molecular simulation in the materials discovery process but also underscores how even small chemical changes, which can be easily done in simulation, can lead to profound differences in performance. To follow this study, Chapter 3 examines a broader design space of side chains which are based on the high performing chemistry found in Chapter 2. We uncover design rules for the crystalline and amorphous phases which are almost counter to one another which stems from the difference in percolation between the two mesophases.

In Chapters 4-6 we study Diamond Liquid Crystalline Elastomers (D-LCE). Previous studies examined defect-free and entanglement-free networks so we evaluate how the tensile response of these systems is affected by the presence of defects and entanglements in the network topology(Chapter 4). Using simulations to mimic a synthesis process developed in earlier work by an experimental group, we find that the synthesized networks generally show that even with a realistic network topology the mechanical response still shows the

characteristic saw-tooth, with the added caveat that entanglements will disrupt this behavior and cause the stress to diverge once a critical strain is reached. We develop a theoretical analysis to understand how synthesis conditions affect the "quality" of the tensile response and find it to be applicable to many chain lengths and persistence lengths.

Chapters 5, and 6 seek to alter the original chain chemistry used in chapter 4. Through alterations to block architecture (Chapter 5) unique tensile responses were found. Two examples include an increase in the number of ordered domain formations per unit deformation and an isotropization of an inherently anisotropic material. By altering the immiscibility between blocks (Chapter 6) a strain-induced crystallization is observed which not only drastically increases the toughness of the material, but also fundamentally changes how domain deformation occurs.

Many times, it is of interest to reinsert chemical detail into a coarse-grained(CG) model, however for some studies (Chapters 4-6) the actual chemistry the CG model is meant to represent is unknown a priori. For a given atomistic chemistry and functional form of the CG model there exists one optimal set of model parameters such that the CG model best represents the physics of the target chemistry. In our case, we are Inverse Coarse Graining (ICG) which is an undefined process as many chemistries can map onto the same model. To tackle this issue, we have introduced a methodology which takes cues from the relative entropy framework.

In chapter 7 we develop a framework to tackle the problem of Inverse Coarse Graining, i. e. the process of finding a chemistry for which a chemically "identity-less" CG model could represent. A general framework is established which takes candidate molecules, maps them to the CG model of interest, and evaluates how well the candidate is "represented" by the CG model. Those which are well represented can be modified in some fashion to create an iterative process where the candidates are continually modified until a desirable degree of representation is achieved. We find this methodology can be modified depending on the ICG problem at hand and shows good accuracy in finding the "correct" candidate when the chemical identity of the CG model is already known.

In chapter 8 we develop a Thermodynamic Integration method to calculate the free energy of polymer self-assembly. We implement a guiding field which makes it more energetically favorable for a certain species to be in specific regions, which allows for the assembly of athermal species into segregated domains as a second-order transition, which overcomes the limitation of first-order transitions where the free energy of the spontaneous segregation is unknown. We implement this methodology to locate phase boundaries of rod-coil block copolymers and to give numerical validation for the change in self-assembly behavior seen in BOLA molecules when the architecture of the side chain is altered from linear to swallow-tail.

Looking into the future of our studies on polythiophene derivatices the immediate work involves working with experimental collaborators to materialize some of the high performing chemistries found in Chapter 3. Agreement between our results and the results

from experiments would further validate our models and give us confidence to extend into a broader design space.

To navigate a much broader design space, we see machine learning as a promising tool to aid materials design. Both the Genetic Algorithm (GA)[121] and the neural network (NN)[181,212] approaches have been employed in design problems that have very large design spaces. To effectively employ these techniques, the development of a platform to automate the process of performing all the simulations is required to calculate a performance score as these techniques require either an iterative process of evaluating multiple generations of candidates (for GA) or an initial evaluation of a large number of candidates (NN)[32,142]. This would be most suitable for cases when the chains organize in a regular arrangement like in the crystalline domain since those positions are predetermined, allowing for the automated placement of the chains. Additionally, the chain design approach would also require careful consideration so that only cases where parameters for all the bonded and nonbonded interactions already exist. Derivation of unique force-field parameters for each chain design would significantly slow down the discovery process, taking away one key advantage of the simulations. Another aspect which needs to be considered is how the chemistry is represented to the algorithms. Many methods have been used in transforming small molecules into machine readable “features” and no universal method has been established. This is because the description can range from a simple list of what functional groups exist in the chemistry, all the way to a complex deconstruction of the bonding topology into a string. If our chain design is only PEO segments and varies in length only, then a simple number describing the number of

segments is suitable. If we wish to include complex chemical motifs or look at non-linear side chain chemistries, then a more complex description is required.

Aside from the side chain chemistry, the semi-crystalline mesophase also has a profound impact on ionic conductivity as well. Preliminary results show that when the orientation of a crystallite is misaligned with the electric field the ionic conductivity through that crystallite shows a sharp decrease. In contrast, the amorphous is found to have no orientational dependence. Experimental results show that the orientation of the crystallites is random which means that the ionic conductivity is lower than it could be if the crystallites were aligned in one direction.

As such, we would like to understand how the overall orientation, size, and number density of the crystallites will affect the effective ionic conductivity. To understand this problem we envision using a Kinetic Monte Carlo (KMC) approach where a semi-crystalline mesophase is represented as a grid of points and the KMC engine uses our results from atomistic simulations to inform the transition rates of the moves the ions can make. This approach will allow us uncork the bunghole and allow for a cascade of new artificially prepared mesophases leading to an understanding of the ionic conduction across the parameter space (crystal size, orientation, and number density) and reveal the underlying trends which can provide a guide to experimental collaborators on what kinds of mesophases to achieve. This study will offer a more complete understanding of ionic conductivity through polythiophene derivatives.

The biggest challenge facing D-LCE currently is finding a chemistry which should reproduce the behaviors seen with the CG models. This is where the use of our ICG

methodology can be leveraged and improved. While the framework we provide in Chapter 7 was successful in identifying good candidates for a given "identity-less" CG model, it is general enough to accommodate new objective functions and methods of generating new candidates. In the context of searching for a D-LCE material one parameter which could be tested for is persistence length, a key measure of stiffness which is a crucial attribute required in the chains to achieve the sought-after saw-tooth tensile response. Additionally, the strength of segregation will be of prime concern when searching for chemistries which can be evaluated by either the radial distribution function between blocks as we used before, or it can be evaluated using a density profile of the two blocks.

Aside from different ways to measure how well a given candidate fits onto the CG model, improvements can be made with respect to how the mapping from the candidate to the CG model is generated. By leveraging a database of literature-extracted data, or a library of structural motifs (fragment library) collected from other users, the correspondence algorithm can be used to try to fit these predefined motifs onto candidate molecules. This would take advantage of atom groupings and mappings already known to work well, giving initial mappings which require fewer refinements. Looking more broadly, recent work has used a graph-theory based approach to find the mapping from a atomistic molecule to a CG model. This generated CG model is not predefined so this approach will not necessarily lead to a mapping which satisfies the constraints of the CG model of interest but it can be used as a method to screen out certain candidates. If this graph theory approach shows that the mapping of a given candidate molecule leads to a CG model which doesn't align with the CG

model of interest then that candidate can be discarded. This can relieve the computational load of running simulations on chemistries which are “doomed” from the outset.

Bibliography

[1] P. Abbott, Math. J. **9**, 689 (2005).

[2] M. Adachi, A. Okumura, E. Sivaniah, and T. Hashimoto, Macromolecules **39**, 7352 (2006).

[3] J. M. Adams and M. Warner, Eur. Phys. J. E **16**, 97 (2005).

[4] B. M. Aguilera-mercado, C. Cohen, and F. A. Escobedo, Macromolecules **47**, 840 (2014).

[5] B. M. Aguilera-mercado, (2012).

[6] K. Alim and E. Frey, Phys. Rev. Lett. **99**, 198102 (2007).

[7] F. H. Arnold, Acc. Chem. Res. **31**, 125 (1998).

[8] M. Asai, T. Katashima, U. Chung, T. Sakai, and M. Shibayama, Macromolecules **46**, 9772 (2013).

[9] T. Asakura and T. Miller, editors , *Biotechnology of Silk* (Springer, New York, n.d.).

[10] J. L. Baker, L. H. Jimison, S. Mannsfeld, S. Volkman, S. Yin, V. Subramanian, A. Salleo, A. P. Alivisatos, and M. F. Toney, Langmuir **26**, 9146 (2010).

[11] A. Bandyopadhyay, P. K. Valavala, T. C. Clancy, K. E. Wise, and G. M. Odegard, Polymer (Guildf). **52**, 2445 (2011).

[12] A. Barducci, M. Bonomi, and M. Parrinello, Wiley Interdiscip. Rev. Comput. Mol. Sci. **1**, 826 (2011).

[13] A. Barducci, G. Bussi, and M. Parrinello, Phys. Rev. Lett. **100**, 020603 (2008).

[14] H. Bässler, Phys. Status Solidi **175**, 15 (1993).

[15] J. R. Beeler Jr., Phys. Rev. **150**, 470 (1966).

[16] C. Berthier, W. Gorecki, M. Minier, M. B. Armand, J. M. Chabagno, and P. Rigaud, Solid State Ionics **11**, 91 (1983).

[17] D. M. Bhawe, C. Cohen, and F. A. Escobedo, Macromolecules **93**, 257804 (2004).

[18] D. M. Bhawe, C. Cohen, and F. A. Escobedo, Macromolecules **37**, 3924 (2004).

[19] J. S. Biggins, M. Warner, and K. Bhattacharya, Phys. Rev. Lett. **103**, 037802 (2009).

[20] E. S. Boek and P. Van Der Schoot, Int. J. Mod. Phys. C **9**, 1307 (1998).

[21] O. Borodin and G. D. Smith, Macromolecules **39**, 1620 (2006).

[22] P. Borsenberger and D. Weiss, *Organic Photoreceptors for Xerography* (Marcel Deckker, New York, 1998).

[23] L. Cai, T. E. Kodger, R. E. Guerra, A. F. Pegoraro, M. Rubinstein, and D. A. Weitz, Adv. Mater. **27**, 5132 (2015).

[24] J. Cannon, J. Aronovitz, and P. Goldbart, Phys. J. **1**, 629 (1991).

[25] A. Chaimovich and M. S. Shell, J. Chem. Phys. **134**, 094112 (2011).

[26] A. Chatterjee and D. G. Vlachos, J. Comput. Aided. Mol. Des. **14**, 253 (2007).

[27] H. Chen, O. Engkvist, Y. Wang, M. Olivercona, and T. Blaschke, Drug Discov. Today **23**, 1241 (2019).

[28] Z. Chen, C. Cohen, and F. A. Escobedo, Macromolecules **35**, 3296 (2002).

[29] C. Cheng, H. Hsueh, C. Lai, C. Pan, B. Hwang, C. Hu, and R. Ho, NPG Asia Mater. **7**, e170 (2015).

[30] A. Chremos, A. Nikoubashman, and A. Z. Panagiotopoulos, J. Chem. Phys. **140**, 054909 (2014).

[31] M. Christen and W. F. van Gunsteren, J. Chem. Phys. **124**, 154106 (2013).

[32] J. Clark, I. Koprinska, and J. Poon, Proc. IEEE 702 (2003).

[33] Y. Cohen, R. J. Albalak, B. J. Dair, M. S. Capel, and E. L. Thomas, Macromolecules **33**, 6502 (2000).

[34] D. Corbett and M. Warner, Sensors Actuators A; Phys. **149**, 120 (2009).

[35] A. J. Crane, F. J. Martinez-veracoechea, F. A. Escobedo, and E. A. Müller, Soft Matter **4**, 1820 (2008).

[36] E. J. W. Crossland, M. Kamperman, M. Nedelcu, C. Ducati, U. Wiesner, D. Smilgies, G. E. S. Toombes, M. A. Hillmyer, S. Ludwigs, U. Steiner, and H. J. Snaith, Nano Lett. **9**, 2807 (2008).

[37] J. Dacuńa and A. Salleo, Phys. Rev. B **84**, 195209 (2011).

[38] P. Dalhaimer, D. E. Discher, and T. C. Lubensky, Nat. Phys. **3**, 354 (2007).

[39] J. J. de Pablo, B. Jones, C. L. Kovacs, V. Ozolins, and A. P. Ramirez, Curr. Opin. Solid Stat Mater. Sci. **18**, 99 (2019).

[40] C. Delacourt, L. Laffont, R. Bouchet, C. Wurm, J.-B. Leriche, M. Morcrette, J.-M. Tarascon, and C. Masquelier, J. Electrochem. Soc. **152**, A913 (2005).

[41] M. Delaye, R. Ribotta, and G. Durand, Phys. Lett. **44**, 139 (1973).

[42] C. Dellago, P. G. Bolhuis, F. S. Csajka, and D. Chandler, J. Chem. Phys. **108**, 1964 (1998).

[43] D. M. Delongchamp, R. J. Kline, D. A. Fischer, L. J. Richter, and M. F. Toney, Adv. Mater. **23**, 319 (2011).

[44] M. Dijkstra and D. Frenkel, Phys. Rev. E **51**, 5891 (1995).

[45] B. X. Dong, B. Huang, A. Tan, and P. F. Green, J. Phys. Chem. C **118**, 17490 (2014).

[46] B. X. Dong, C. Nowak, J. W. Onorato, J. Strzalka, F. A. Escobedo, C. K. Luscombe, P. F. Nealey, and S. N. Patel, Chem. Mater. **118**, 1418 (2019).

[47] H. Dong, X. Fu, J. Liu, Z. Wang, and W. Hu, Adv. Mater. **25**, 6158 (2013).

[48] J. P. K. Doye, R. P. Sear, and D. Frenkel, J. Chem. Phys. **208**, 2134 (1998).

[49] L. G. Dunfield, A. W. Burgess, and H. A. Scheraga, J. Phys. Chem. **82**, 2609 (1978).

[50] D. T. Duong, V. Ho, Z. Shang, S. Mollinger, S. C. B. Mannsfeld, J. Dacuña, M. F. Toney, R. Segalman, and A. Salleo, Adv. Funct. Mater. **24**, 4515 (2014).

[51] D. J. Earl and M. W. Deem, Phys. Chem. Chem. Phys. **7**, 3910 (2005).

[52] S. F. Edwards and T. Vilgis, Polymer (Guildf). **27**, 483 (1986).

[53] F. Ercolessi and J. B. Adams, Europhys. Lett. **26**, 583 (1994).

[54] F. A. Escobedo, E. E. Borrero, and J. C. Araque, J. Phys. Condens. Matter **21**, 333101 (2009).

[55] F. A. Escobedo and J. J. de Pablo, J. Chem. Phys. **106**, 9858 (1998).

[56] P. Español and P. B. Warren, Europhys. Lett. **30**, 191 (1995).

[57] R. Faller, F. Müller-Plathe, M. Doxastakis, and D. Theodorou, Macromolecules **34**, 1436 (2001).

[58] R. Fox, J. Theor. Biol. **234**, 187 (2005).

[59] P. L. Freddolino, A. S. Arkhipov, S. B. Larson, A. Mcpherson, and K. Schulten, Structure **14**, 437 (2006).

[60] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Academic Press, New York, 2002).

[61] H. Funabashi, Y. Miyamoto, Y. Isono, T. Fujimoto, Y. Matsushita, and N. Mitsuru, Macromolecules **16**, 1 (1983).

[62] R. M. Fuoss and C. A. Kraus, J. Am. Chem. Soc. **55**, 1019 (1933).

[63] A. J. Gabert, E. Verploegen, P. T. Hammond, and R. R. Schrock, Macromolecules **39**, 3993 (2006).

[64] M. L. Gardel, J. H. Shin, F. C. MacKintosh, L. Mahadevan, P. Matsudaira, and D. A. Weitz, Science (80-.). **304**, 1301 (2004).

[65] G. Gaucher, M.-H. Dufresne, V. P. Sant, N. Kang, D. Maysinger, and J.-C. Leroux, J. Control. Release **109**, 169 (2005).

[66] B. Geffroy, P. le Roy, and C. Prat, Polym. Int. **55**, 572 (2006).

[67] A. Ghanbari, M. C. Böhm, and F. Müller-Plathe, Macromolecules **44**, 5520 (2011).

[68] A. Giovannitti, D. Sbircea, S. Inal, C. B. Nielsen, E. Bandiello, D. A. Hanifi, M. Sessolo, G. G. Malliaras, I. Mcculloch, and J. Rivnay, Proc. Natl. Acad. Sci. **113**, 12017 (2016).

[69] R. Giridharagopal, L. Q. Flagg, J. S. Harrison, M. E. Ziffer, J. W. Onorato, C. K. Luscombe, and D. S. Ginger, Nat. Mater. **16**, 737 (2017).

[70] S. C. Glotzer and W. Paul, Annu. Rev. Mater. Res. **32**, 401 (2002).

[71] G. S. Grest and K. Kremer, Phys. Rev. A **33**, 3628 (1986).

[72] G. S. Grest, M.-D. Lacasse, K. Kremer, and A. M. Gupta, J. Chem. Phys. **105**, 10583 (1996).

[73] G. Grochola, J. Chem. Phys. **120**, 2122 (2004).

[74] R. D. Groot and P. B. Warren, J. Chem. Phys. **107**, 4423 (1997).

[75] F. Grunewald, G. Rossi, A. H. de Vries, S. J. Marrink, and L. Monticelli, J. Phys. Chem. B **122**, 7436 (2018).

[76] S. Günes, H. Neugebauer, and N. S. Sariciftci, Chem. Rev. **107**, 1324 (2007).

[77] C. S. Haines, M. D. Lima, N. Li, G. M. Spinks, J. Foroughi, J. D. W. Madden, S. H. Kim, S. Fang, M. J. de Andrade, F. Göktepe, Ö. Göktepe, S. M. Mirvakili, S. Naficy, X. Lepró, J. Oh, M. E. Kozlov, S. J. Kim, X. Xu, B. J. Swedlove, G. G. Wallace, and R. H. Baughman, Science (80-.). **343**, 868 (2014).

[78] D. A. Hajduk, P. E. Harper, S. M. Gruner, C. C. Honeker, G. Kim, E. L. Thomas, and L. J. Fetters, Macromolecules **27**, 4063 (1994).

[79] K. Hansen, N. Dau, F. Feist, C. Deck, R. Willinger, S. M. Madey, and M. Bottlang, Accid. Anal. Prev. **59**, 109 (2013).

[80] V. A. Harmandaris, D. Reith, N. F. A. van der Vegt, and K. Kremer, Macromol. Chem. Phys. **208**, 2109 (2007).

[81] S. Himmelberger, J. Dacuña, J. Rivnay, L. H. Jimison, T. Mccarthy-ward, M. Heeney, I. Mcculloch, M. F. Toney, and A. Salleo, Adv. Funct. Mater. **23**, 2091 (2013).

[82] M. A. Horsch, Z. Zhang, and S. C. Glotzer, J. Chem. Phys. **125**, 184903 (2006).

[83] M. A. Horsch, Z. Zhang, C. R. Iacovella, and S. C. Glotzer, J. Chem. Phys. **121**, 11455 (2004).

[84] C. N. Hoth, P. Schilinsky, S. A. Choulis, and C. J. Brabec, Nano Lett. **8**, 2806 (2008).

[85] H. Hsueh, Y. Huang, R. Ho, C. Lai, T. Makida, and H. Hasegawa, Adv. Mater. **23**, 3041 (2011).

[86] D. M. Huang, R. Faller, K. Do, and A. J. Moulé, J. Chem. Theory Comput. **6**, 526 (2010).

[87] W. Humphrey, A. Dalke, and K. Schulten, J. Mol. Graph. **14**, 33 (1996).

[88] S. Hur, V. Thapar, A. Ramírez-hernández, G. Khaira, T. Segal-peretz, P. A. Rincon-Delgadillo, W. Li, M. Müller, P. F. Nealey, and J. J. de Pablo, Proc. Natl. Acad. Sci. **112**, 14144 (2015).

[89] S. Inal, G. G. Malliaras, and J. Rivnay, J. Mater. Chem. C **4**, 3942 (2016).

[90] S. Inal, J. Rivnay, A.-O. Suiu, G. G. Malliaras, and I. Mcculloch, Acc. Chem. Res. **19**, 3201 (2018).

[91] K. C. Ingham, S. A. Brew, S. Huff, and S. V Litvinovich, J. Biol. Chem. **272**, 1718 (1997).

[92] H. I. Ingólfsson, C. A. Lopez, J. J. Uusitalo, D. H. De Jong, S. M. Gopal, X. Periole, and S. J. Marrink, Adv. Rev. **4**, 225 (2014).

[93] S. Izvekov and G. A. Voth, J. Phys. Chem. B **109**, 2469 (2005).

[94] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, Apl Mater. **1**, 011002 (2013).

[95] P. A. Janmey, M. E. McCormick, S. Rammensee, J. L. Leight, P. C. Georges, and F. C. MacKintosh, Nat. Mater. **6**, 48 (2007).

[96] A. E. Javier, S. N. Patel, D. T. Hallinan Jr, V. Srinivasan, and N. P. Balsara, Angew. Chemie Int. Ed. **50**, 9848 (2011).

[97] L. H. Jimison, S. Himmelberger, D. T. Duong, J. Rivnay, M. F. Toney, and A. Salleo, J. Polym. Sci. Part B Polym. Phys. **51**, 611 (2013).

[98] D. H. De Jong, G. Singh, W. F. D. Bennett, C. Arnarez, T. A. Wassenaar, L. V Schäfer, X. Periole, D. P. Tieleman, and S. J. Marrink, J. Chem. Theory Comput. **9**, 687 (2012).

[99] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, J. Am. Chem. Soc. **118**, 11225 (1996).

[100] A. Jung, L. A. A. Beex, S. Diebels, and S. P. A. Bordas, Mater. Des. **87**, 36 (2015).

[101] H. Kamberaj, R. J. Low, and M. P. Neal, J. Chem. Phys. **122**, 244114 (2005).

[102] H. A. Karimi-varzaneh, N. F. A. van der Vegt, F. Müller-plathe, and P. Carbone, ChemPhysChem **13**, 3428 (2012).

[103] N. Kayunkid, S. Uttiya, and M. Brinkmann, Macromolecules **43**, 4961 (2010).

[104] D. Khodagholy, T. Doublet, P. Quilichini, M. Gurfinkel, P. Leleux, A. Ghestem, E. Ismailova, T. Herve, S. Sanaur, C. Bernard, and G. G. Malliaras, Nat. Commun. **4**, 1575 (2013).

[105] F. Kienberger, V. P. Pastushenko, G. Kada, H. J. Gruber, C. Riener, H. Schindler, and P. Hinterdorfer, Single Mol. **1**, 123 (2000).

[106] H. Kim, J. M. Boothby, S. Ramachandran, C. D. Lee, and T. H. Ware, Macromolecules **50**, 4267 (2017).

[107] A. Kitao, K. Yonekura, S. Maki-yonekura, F. A. Samatey, K. Imada, K. Namba, and N. Go, Proc. Natl. Acad. Sci. **103**, 4894 (2006).

[108] R. J. Kline, M. D. Mcgehee, E. N. Kadnikova, J. Liu, J. M. J. Fréchet, and M. F. Toney, Macromolecules **38**, 3312 (2005).

[109] E. Klotzsch, M. L. Smith, K. E. Kubow, S. Muntwyler, W. C. Little, F. Beyeler, D. Gourdon, B. J. Nelson, and V. Vogel, Proc. Natl. Acad. Sci. **106**, 18267 (2009).

[110] D. P. Knight and F. Vollrath, Philos. Trans. R. Soc. London. Ser. B Biol. Sci. **357**, 155 (2002).

[111] K. Koo, H. Ahn, S. Kim, D. Ryu, and T. P. Russell, Soft Matter **9**, 9059 (2013).

[112] K. Kremer and G. S. Grest, J. Chem. Phys. **92**, 5057 (1990).

[113] R. Kroon, D. Kiefer, D. Stegerer, L. Yu, M. Sommer, and C. Müller, Adv. Mater. **29**, 1700930 (2017).

[114] R. Kroon, M. Lenes, J. C. Hummelen, P. W. M. Blom, and B. De Boer, Polym. Rev. **48**, 531 (2008).

[115] O. Kuchner and F. H. Arnold, Trends Biotechnol. **15**, 523 (1997).

[116] S. Kullback and R. A. Leibler, Ann. Math. Stat. **22**, 79 (2019).

[117] J. Küpfer and H. Finkelmann, Die Makromol. Chemie, Rapid Commun. **12**, 717 (1991).

[118] S. Lafon and A. B. Lee, IEEE Trans. Pattern Anal. Mach. Intell. **28**, 1393 (2006).

[119] C. Lai, C. Ho, H. Chen, and W. Su, Macromolecules **46**, 2249 (2013).

[120] A. Laio and M. Parrinello, Proc. Natl. Acad. Sci. **99**, 12562 (2002).

[121] R. Le Riche and R. T. Haftkat, AIAA J. **31**, 951 (1993).

[122] A. B. Lee and B. Nadler, Artif. Intell. Stat. 259 (2007).

[123] K. Lee, L. K. Povlich, and J. Kim, Analyst **135**, 2179 (2010).

[124] M. Li, P. Keller, B. Li, X. Wang, and M. Brunet, Adv. Mater. **15**, 569 (2003).

[125] M. Li, P. Keller, J. Yang, and P.-A. Albouy, Adv. Mater. **16**, 1922 (2004).

[126] T. Liu, L. Finn, M. Yu, H. Wang, T. Zhai, X. Lu, Y. Tong, and Y. Li, Nano Lett. **14**, 2522 (2014).

[127] X. J. Loh and O. A. Scherman, *Polymeric and Self Assembled Hydrogels: From Fundamental Understanding to Applications* (Royal Society of Chemistry, London, 2012).

[128] F. J. Martinez-veracoechea and F. A. Escobedo, Macromolecules **40**, 7354 (2007).

[129] F. J. Martínez-veracoechea and F. A. Escobedo, J. Chem. Phys. **125**, 104907 (2006).

[130] M. W. Matsen and F. S. Bates, Macromolecules **29**, 7641 (1996).

[131] M. W. Matsen and M. Schick, Phys. Rev. Lett. **72**, 2660 (1994).

[132] L. R. Middleton, S. Szewczyk, J. Azoulay, D. Murtagh, G. Rojas, K. B. Wagener, J. Cordaro, and K. I. Winey, Macromolecules **48**, 3713 (2015).

[133] M. Misra, M. Agarwal, D. W. Sinkovits, S. K. Kumar, C. Wang, G. Pilania, R. Ramprasad, R. A. Weiss, X. Yuan, and T. C. M. Chung, Macromolecules **47**, 1122 (2014).

[134] M. Misra and S. K. Kumar, ACS Macro Lett. **6**, 200 (2017).

[135] S. A. Mollinger, B. A. Krajina, R. Noriega, A. Salleo, and A. J. Spakowitz, ACS Macro Lett. **4**, 708 (2015).

[136] T. C. Moore, C. R. Iacovella, and C. Mccabe, J. Chem. Phys. **140**, 06B606_1 (2014).

[137] K. Morishige and N. Tarui, J. Phys. Chem. **111**, 280 (2007).

[138] H. Morita, J. Photopolym. Sci. Technol. **26**, 801 (2013).

[139] R. J. Mortimer, A. L. Dyer, and J. R. Reynolds, Displays **27**, 2 (2006).

[140] K. Müllen and U. Scherf, *Organic Light Emitting Devices: Synthesis, Properties and Applications* (CRC Press, New York, 2006).

[141] M. Müller and K. C. Daoulas, J. Chem. Phys. **128**, 024903 (2008).

[142] N. Nguyen and A. Cripps, J. Real Estate Res. **22**, 313 (2001).

[143] C. B. Nielsen, A. Giovannitti, D. Sbircea, E. Bandiello, M. R. Niazi, D. A. Hani, M. Sessolo, A. Amassian, G. G. Malliaras, J. Rivnay, and I. Mcculloch, J. Am. Chem. Soc. **138**, 10252 (2016).

[144] K. Nishi, K. Fujii, Y. Katsumoto, T. Sakai, and M. Shibayama, Macromolecules **47**, 3274 (2014).

[145] W. G. Noid, J. Chu, G. S. Ayton, V. Krishna, S. Izvekov, G. A. Voth, A. Das, and H. C. Andersen, J. Chem. Phys. **128**, 244114 (2008).

[146] R. Noriega, J. Rivnay, K. Vandewal, F. P. V Koch, N. Stingelin, P. Smith, M. F. Toney, and A. Salleo, Nat. Mater. **12**, 1038 (2013).

[147] S. Nosé, J. Chem. Phys. **81**, 511 (1984).

[148] A. Nova, S. Keten, N. M. Pugno, A. Redaelli, and M. J. Buehler, Nano Lett. **10**, 2626 (2010).

[149] C. Nowak and F. A. Escobedo, Macromolecules **49**, 6711 (2016).

[150] C. Nowak and F. A. Escobedo, J. Chem. Theory Comput. **14**, 5984 (2018).

[151] C. Nowak and F. A. Escobedo, Macromolecules **51**, 5685 (2018).

[152] C. Nowak and F. A. Escobedo, Phys. Rev. Mater. **1**, 035601 (2017).

[153] P. Olmsted, Phys. J. **4**, 2215 (1994).

[154] R. M. Owens and G. G. Malliaras, MRS Bull. **35**, 449 (2019).

[155] P. Padmanabhan, M. Chavis, C. K. Ober, and F. A. Escobedo, Soft Matter **10**, 6172 (2014).

[156] P. Padmanabhan, F. J. Martinez-veracoechea, J. C. Araque, and F. A. Escobedo, J. Chem. Phys. **136**, 234905 (2012).

[157] P. Padmanabhan, F. Martinez-veracoechea, and F. A. Escobedo, Macromolecules **49**, 5232 (2016).

[158] S. N. Patel, A. E. Javier, G. M. Stone, S. A. Mullin, and N. P. Balsara, ACS Nano **6**, 1589 (2012).

[159] Z. Pei, Y. Yang, Q. Chen, E. M. Terentjev, Y. Wei, and Y. Ji, Nat. Mater. **13**, 36 (2013).

[160] D. M. Pesko, Y. Jung, A. L. Hasan, M. A. Webb, G. W. Coates, T. F. Miller III, and N. P. Balsara, Solid State Ionics **289**, 118 (2016).

[161] D. M. Pesko, M. A. Webb, Y. Jung, Q. Zheng, T. F. Miller, G. W. Coates, and N. P. Balsara, Macromolecules **49**, 5244 (2016).

[162] A. Pipertzis, M. Mu, M. Mezger, U. Scherf, and G. Floudas, Macromolecules **51**, 6440 (2018).

[163] S. Plimpton, J. Comput. Phys. **117**, 1 (1995).

[164] C. Poelking and D. Andrienko, Macromolecules **46**, 8941 (2013).

[165] M. Pütz, K. Kremer, and G. S. Grest, Europhys. Lett. **49**, 735 (2000).

[166] R. Ramachandran, G. Beaucage, A. S. Kulkarni, D. Mcfaddin, J. Merrick-mack, and V. Galiatsatos, Macromolecules **41**, 9802 (2008).

[167] B. O. Regan and M. Gratzelt, Lett. to Nat. **353**, 737 (1991).

[168] H. Reich, M. Dijkstra, R. van Roji, and M. Schmidt, J. Phys. Chem. B **111**, 7825 (2007).

[169] D. Reith, M. Pütz, and F. Müller-Plathe, J. Comput. Chem. **24**, 1624 (2003).

[170] Y. M. Rhee and V. S. Pande, Biophys. J. **84**, 775 (2003).

[171] M. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez, and H. E. Gaub, Science (80-.). **276**, 1109 (1997).

[172] J. Rivnay, S. Inal, B. A. Collins, M. Sessolo, E. Stavrinidou, X. Strakosas, C. Tassone, D. M. Delongchamp, and G. G. Malliaras, Nat. Commun. **7**, 11287 (2016).

[173] J. Rivnay, R. M. Owens, and G. G. Malliaras, Chem. Mater. **26**, 679 (2014).

[174] V. Rühle, C. Junghans, A. Lukyanov, K. Kremer, and D. Andrienko, J. Chem. Theory Comput. **5**, 3211 (2009).

[175] T. Sakai, T. Matsunaga, Y. Yamamoto, C. Ito, R. Yoshida, S. Suzuki, N. Sasaki, M. Shibayama, and U. Chung, Macromolecules **41**, 5379 (2008).

[176] G. Santangelo, A. Di Matteo, F. Müller-Plathe, and G. Milano, J. Phys. Chem. B **111**, 2765 (2007).

[177] N. Sary, C. Brochon, G. Hadziioannou, and R. Mezzenga, Eur. Phys. J. E **24**, 379 (2007).

[178] E. S. Savoy and F. A. Escobedo, Langmuir **28**, 16080 (2012).

[179] M. R. J. Scherer, P. M. S. Cunha, and U. Steiner, Adv. Mater. **26**, 2403 (2014).

[180] D. A. Scherson and A. Palencsár, Interface **15**, 17 (2006).

[181] J. Schmidhuber, Neural Networks **61**, 85 (2015).

[182] A. J. Schultz, C. K. Hall, and J. Genzer, Macromolecules **40**, 2629 (2007).

[183] A. J. Schultz, C. K. Hall, and J. Genzer, J. Chem. Phys. **120**, 2049 (2004).

[184] K. Schwenke, M. Lang, and J. U. Sommer, Macromolecules **44**, 9464 (2011).

[185] T. Sekitani, U. Zschieschang, H. Klauk, and T. Someya, Nat. Mater. **9**, 1015 (2010).

[186] S. Sen, S. K. Kumar, and P. Keblinski, Macromolecules **38**, 650 (2005).

[187] M. S. Shell, J. Chem. Phys. **129**, 144108 (2008).

[188] G. Sigaud, D. Y. Yoon, and A. C. Griffin, Macromolecules **16**, 875 (1983).

[189] H. Sirringhaus, Adv. Mater. **26**, 1319 (2014).

[190] E. J. Smith, T. Bryk, and A. D. J. Haymet, J. Chem. Phys. **123**, 034706 (2005).

[191] P. E. Smith, R. M. Brame, A. E. Mark, and W. F. van Gunsteren, J. Phys. Chem. **97**, 2009 (2014).

[192] T. L. Smith and J. E. Frederick, J. Appl. Phys. **36**, 2996 (1965).

[193] J. G. Son, M. Son, K. Moon, B. H. Lee, J. Myoung, M. S. Strano, M. Ham, and C. A. Ross, Adv. Mater. **25**, 4723 (2013).

[194] A. K. Soper, Chem. Phys. **202**, 295 (1996).

[195] T. Spyriouni, C. Tzoumanekas, D. Theodorou, F. Müller-Plathe, and G. Milano, Macromolecules **40**, 3876 (2007).

[196] E. Stavrinidou, P. Leleux, H. Rajaona, D. Khodagholy, J. Rivnay, M. Lindau, S. Sanaur, and G. G. Malliaras, Adv. Mater. **25**, 4488 (2013).

[197] A. Sugimura, M. Asai, T. Matsunaga, Y. Akagi, T. Sakai, H. Noguchi, and M. Shibayama, Polym. J. **45**, 300 (2012).

[198] Y. Sugita and Y. Okamoto, Chem. Phys. Lett. **314**, 141 (1999).

[199] S. K. Sukumaran, G. S. Grest, K. Kremer, and R. Everaers, J. Polym. Sci. Part B Polym. Phys. **43**, 917 (2005).

[200] S. Sun, Sol. Energy Mater. Sol. Cells **79**, 257 (2003).

[201] Y. Sun, P. Padmanabhan, M. Misra, and F. A. Escobedo, Soft Matter **13**, 8542 (2017).

[202] C. Svaneborg, H. A. Karimi-varzaneh, N. Hojdis, F. Fleck, and R. Everaers, ArXiv 1 (2018).

[203] R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **57**, 2607 (1986).

[204] A. R. Tajbakhsh and E. M. Terentjev, Eur. Phys. J. E **6**, 181 (2001).

[205] J. Tang, Y. Jiang, X. Zhang, D. Yan, and J. Z. Y. Chen, Macromolecules **48**, 9060 (2015).

[206] V. Thapar and F. A. Escobedo, Phys. Rev. Lett. **112**, 048301 (2014).

[207] T. T. To and S. Adams, Phys. Chem. Chem. Phys. **16**, 4653 (2014).

[208] C. Tschierske, Chem. Soc. Rev. **36**, 1930 (2007).

[209] K. Urayama, Macromolecules **40**, 2277 (2007).

[210] K. Urayama, T. Kawamura, and S. Kohjiya, Polymer (Guildf). **50**, 347 (2008).

[211] K. Urayama, E. Kohmon, M. Kojima, and T. Takigawa, Macromolecules **42**, 4084 (2009).

[212] M. van Gereven and S. Bhte, *Artificial Neural Networks as Models of Neural Information Processing* (Frontiers Media SA, 2018).

[213] P. Vashishta, R. K. Kalia, and A. Nakano, J. Phys. Chem. B **110**, 3727 (2006).

[214] A. M. Veselinovic, J. B. Veselinovic, J. V. Zivkovic, and G. M. Nikolic, Curr. Top. Med. Chem. **15**, 1768 (2015).

[215] T. Vettorel, A. Y. Grosberg, and K. Kremer, Phys. Biol. **6**, 025013 (2009).

[216] D. G. Walton, G. J. Kellogg, A. M. Mayes, P. Lambooy, and T. P. Russell, Macromolecules **27**, 6225 (1994).

[217] E. Wang and F. A. Escobedo, Macromolecules **49**, 2375 (2016).

[218] T. H. Ware, M. E. McConney, J. J. Wie, V. P. Tondiglia, and T. J. White, Science (80-.). **347**, 982 (2015).

[219] M. Warner and E. M. Terentjev, *Liquid Crystal Elastomers* (Oxford University Press, Oxford, 2007).

[220] T. A. Wassenaar, K. Pluhackova, R. A. Böckmann, S. J. Marrink, and D. P. Tieleman, J. Chem. Theory Comput. **10**, 676 (2013).

[221] D. Weininger, J. Chem. Inf. Comput. Sci. **28**, 31 (1988).

[222] H. Wermter and H. Finkelmann, E-Polymers **1**, (2001).

[223] D. R. Wheeler and J. Newman, J. Chem. Phys. **108**, 18353 (2004).

[224] D. Whitley, Statustics Comput. **4**, 65 (1994).

[225] J. Wilhelm and E. Frey, Phys. Rev. Lett. **91**, 108103 (2003).

[226] D. R. M. Williams and M. Warner, Phys. J. **51**, 317 (1990).

[227] M. C. Wintersgill, J. J. Fontanella, Y. S. Pak, S. G. Greenbaum, A. Al-Mudaris, and A. V. Chadwick, Polymer (Guildf). **30**, 1123 (1989).

[228] X. Wu, X. Liu, N. Du, G. Xu, and B. Li, Appl. Phys. Lett. **95**, 093703 (2009).

[229] J. Xia, W. F. Flynn, E. Gallicchio, B. W. Zhang, P. He, Z. Tan, and R. M. Levy, J. Comput. Chem. **36**, 1772 (2015).

[230] Z. Xu and X. Li, Adv. Funct. Mater. **21**, 3883 (2011).

[231] M. Yamanoi, O. Pozo, and J. M. Maia, J. Chem. Phys. **135**, 044904 (2011).

[232] H. Yang, A. Buguin, J. Taulemesse, K. Kaneko, S. Méry, A. Bergeret, and P. Keller, J. Am. Chem. Soc. **131**, 15000 (2009).

[233] P. Yi, C. R. Locker, and G. C. Rutledge, Macromolecules **46**, 4723 (2013).

[234] Y. Yi, X. Fan, X. Wan, L. Li, N. Zhao, X. Chen, J. Xu, and Q. Zhou, Macromolecules **37**, 7610 (2004).

[235] X. Zhang, H. Bronstein, A. J. Kronemeijer, J. Smith, Y. Kim, R. J. Kline, L. J. Richter, T. D. Anthopoulos, H. Sirringhaus, K. Song, M. Heeney, W. Zhang, I. Mcculloch, and D. M. Delongchamp, Nat. Commun. **4**, 2238 (2013).

[236] J. Ziv, IEEE Trans. Pattern Anal. Mach. Intell. **24**, 530 (1978).

[237] J. Ziv and A. Lempel, IEEE Trans. Pattern Anal. Mach. Intell. **23**, 337 (1977).